

AD-A088 548

BOLT BERANEK AND NEWMAN INC CAMBRIDGE MA  
WAVEFRONT DECONVOLUTION STUDIES. (U)  
MAY 80 R BARAKAT

F/6 20/10

UNCLASSIFIED

RADC-TR-80-154

F30602-79-C-0018

NL

1 OF 1  
AD-A088 548

END  
DATE  
FILMED  
10-80  
DTIC

AD A088548

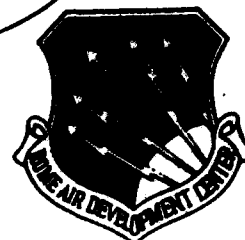
**RADC-TR-80-154**

**Interim Report**

**May 1980**

**LEVEL II**

(12)



# **WAVEFRONT DECONVOLUTION STUDIES**

**Bolt, Beranek & Newman, Inc.**

Sponsored by  
Defense Advanced Research Projects Agency (DoD)  
ARPA Order No. C223

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

**ROME AIR DEVELOPMENT CENTER**  
**Air Force Systems Command**  
**Griffiss Air Force Base, New York 13441**

**DTIC**  
**ELECTE**  
**SEP 2 1980**

**A**

**80 8 29 08**

DOUG FILE 1004

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-80-154 has been reviewed and is approved for publication.

APPROVED:



DORIS HAMILL, 1Lt, USAF  
Project Engineer

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (OCSE), Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return this copy. Retain or destroy.

## WAVEFRONT DECONVOLUTION STUDIES

Dr. Richard Barakat

Contractor: Bolt, Beranek & Newman, Inc.  
Contract Number: F30602-79-C-0018  
Effective Date of Contract: 29 Nov 78  
Contract Expiration Date: 28 Nov 79  
Short Title of Work: Wavefront Deconvolution Studies  
Program Code Number: OE20  
Period of Work Covered: Nov 78 - Nov 79

Principal Investigator: Dr. Richard Barakat  
Phone: (617) 491-1850  
Project Engineer: Doris Hamill, 1Lt, USAF  
Phone: (315) 330-3148

Approved for public release; distribution unlimited.

This research was supported by the Defense Advanced Research Projects Agency of the Department of Defense and was monitored by Lt Doris Hamill (OCSE), Griffiss AFB NY 13441 under Contract F30602-79-C-0018.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-80-154	2. GOVT ACCESSION NO. AD-A088	3. RECIPIENT'S CATALOG NUMBER 548
4. TITLE (and Subtitle) WAVEFRONT DECONVOLUTION STUDIES	5. TYPE OF REPORT & PERIOD COVERED Interim Report Nov 78 - Nov 79	6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) Dr. Richard Barakat	8. CONTRACT OR GRANT NUMBER(s) F30602-79-C-0018	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Bolt, Beranek & Newman, Inc. 50 Moulton Street Cambridge MA 02238	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62301E C2230108	
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd Arlington VA 22209	12. REPORT DATE May 1980	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Rome Air Development Center (OCSE) Griffiss AFB NY 13441	13. NUMBER OF PAGES 90	
	15. SECURITY CLASS. (of this report) UNCLASSIFIED	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) Same		
18. SUPPLEMENTARY NOTES RADC Project Engineer: Doris Hamill, 1Lt, USAF (OCSE)		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Wavefront deconvolution Zernike polynomials Singular value decomposition Wavefront aberrations		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This interim report summarizes the research essentially completed during the report period in question. The first section describes a numerically stable iterative method for the inversion of wavefront aberrations from measured point spread function data. Section two is devoted to the development of generalizations of Zernike polynomials that are useful in studying optimum balanced wavefront aberrations. The third section outlines the method of filtered singular value decomposition to wavefront deconvolution.		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

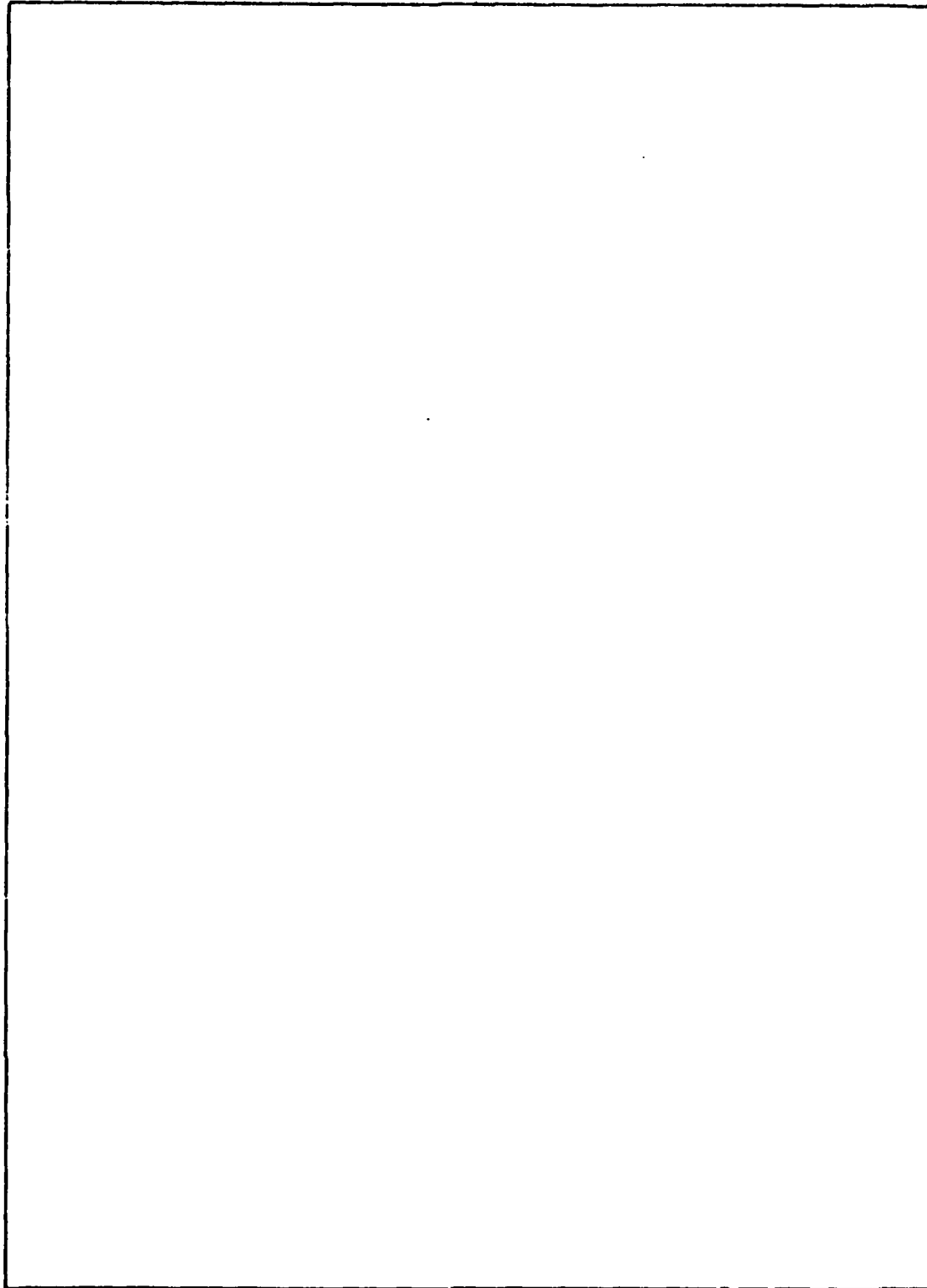
UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

060100 24

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

## GENERAL INTRODUCTION

The present final report chronicles various aspects of the research conducted by the principal investigator related to the wavefront deconvolution problem. The material is summarized in the form of three self-contained sections, each devoted to a single item. The three sections are:

*Section 1:* A numerically stable iterative method for the inversion of wavefront aberrations from measured point spread function data.

*Section 2:* Optimum balanced wavefront aberrations for radially symmetric amplitude distributions; generalizations of Zernike polynomials.

*Section 3:* Application of filtered singular value decomposition to wavefront deconvolution.

Accession For	
NTIS Grant	<input checked="checked" type="checkbox"/>
DDC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution	
Approved by	
Dist	Approved for
A	Special

SECTION 1  
A NUMERICALLY STABLE ITERATIVE METHOD FOR THE  
INVERSION OF WAVEFRONT ABERRATIONS FROM  
MEASURED POINT SPREAD FUNCTION DATA

ABSTRACT

This paper outlines a method for the determination of the unknown wavefront aberration function of an optical system from noisy measurements of the corresponding point spread function. The problem is cast as a nonlinear least squares estimation problem for the values of the wavefront aberration function at  $N$  points over the slit aperture, from measurements of the point spread function at  $M$  points with  $M \geq N$ . Newton's method is used to replace the nonlinear minimization problem with a sequence of linear problems. Each such problem requires the inversion of the Hessian matrix of the error metric which is shown to be both singular (with rank  $\leq N-1$ ) and ill-conditioned. To overcome singularity, the pseudoinverse is used; to overcome ill-conditioning the pseudoinverse is calculated using singular value decomposition and the singular values then filtered. Attention is drawn to difficulties such as nonuniqueness, sensitivity of algorithms to initial guess, etc., the ancillary mathematical details being set out in appendices. Some illustrative numerical results are presented and analyzed.



## 1. INTRODUCTION

There are a number of situations of current interest where one is required to determine the wavefront aberration function of the optical system from measurements of the corresponding point spread function (e.g., image forming adaptive optical systems, laser beam forming, etc.). Published papers *specifically* devoted to this problem are: Gonsalves [1] and Southwell [2]. The general problem of phase retrieval from modulus data, of which this is a typical situation, has been attacked by a variety of methods almost too diverse to catalog and we refer the reader to the vast literature for details. Some typical references are [3-10].

Our approach to the problem rests upon two provisos:

1. The aberrated wavefront itself is the primary artifact of the inversion, not an assumed functional form of it. Curve fitting the reconstructed wavefront can be done *after* the inversion, if desired.

2. The inversion of the wavefront from the measured point spread function involves the solution of a nonlinear integral equation of the first kind, Eq. 2.2, with the attendant numerical instability as befits an ill-posed problem [11,12]. The nonlinear inversion method is therefore tailored to be robust with respect

to noise in the measured point spread function.

Given these two provisos, we have chosen to cast the problem of determining the wavefront as a nonlinear least squares estimation problem and brought to bear the powerful tools of modern numerical analysis towards a solution.

The plan of the paper is as follows. Section 2 is devoted to the necessary preliminary material. In section 3 is discussed the strategy of the nonlinear least squares, while section 4 contains the tactics (filtered singular value decomposition, the scaling conditions, stopping rule, etc.) required for performing the inversion. The ancillary mathematical details are relegated to a series of Appendices. Finally some numerical results are discussed in section 5.

Although our analysis is couched in the specifics of wavefront aberration and point spread function, the formalism is independent of the specific situation and is applicable to the phase retrieval problem in general, for phases having compact support.

## 2. PRELIMINARIES

The measured point spread function will be termed  $\tau_m$  where  $m = 1, 2, \dots, M$  are the indexed values. It will be convenient to write

$$\hat{\tau} \equiv \begin{vmatrix} \tau_1 \\ \tau_2 \\ \cdot \\ \cdot \\ \tau_m \end{vmatrix} \quad (2.1)$$

We take as the diffraction model

$$t(v) = \left| \frac{1}{2} \int_{-1}^{+1} e^{i w(p)} e^{i v p} dp \right|^2 \quad (2.2)$$

where  $w(p) = (2\pi/\lambda)W(p)$  is the wavefront aberration function measured in wavenumber units  $(2\pi/\lambda)$ . Amplitude variations over the exit pupil are not allowed in this version. Although this can be included in the analysis we prefer to omit it in order to focus on what we believe to be the more important issues.

The diffraction model is determined by  $N$  free parameters, namely the values of the wavefront aberration function  $W(p)$  at  $N$  points over the slit aperture. The points  $p_1, p_2, \dots, p_N$  need not be equally spaced. We choose to make them equally spaced and to let  $N$  be an odd integer so as to include the point  $p=0$ . The aberration

function must, by definition, satisfy the requirement

$$W(0) = 0 \quad (2.3)$$

In order to carry out this program, we discretize the integral

$$t(v_m) \equiv t_m = \left| \frac{1}{2} \sum_{n=1}^N \alpha_n e^{iv_m p_n} e^{iw_n} \right|^2 \quad (2.4)$$

Here  $p_n$  are the quadrature points and  $\alpha_n$  the corresponding weight factors. Our calculations were performed using a trapezoidal rule. We have also set  $w_n \equiv w(p_n)$ .

It is again convenient to write the  $N$   $w_n$  parameters as a column vector of length  $N$

$$\hat{W} \equiv \begin{vmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ w_N \end{vmatrix} \quad (2.5)$$

In this condensed notation, Eq. 2.4 becomes

$$t_1(\hat{W}), t_2(\hat{W}), \dots, t_M(\hat{W}) \quad (2.6)$$

indicating that each value of  $t_m$  is related to all the  $w_n$ . In the *direct* problem, we are given  $\hat{W}$  and are required to calculate  $t_m(\hat{W})$ .

The *inverse* (or inversion) problem relates the unknown wave-front values  $w_n$  to the known (measured) point spread function data  $\tau_1, \tau_2, \dots, \tau_M$  via the *nonlinear* functional relation

$$\tau_m = t_m(\hat{W}), m = 1, \dots, M \quad (2.7)$$

This can be set into the more succinct form

$$\hat{\tau} = \hat{t}(\hat{W}) \quad (2.8)$$

upon defining the additional vector

$$\hat{t}(\hat{W}) = \begin{vmatrix} t_1(\hat{W}) \\ t_2(\hat{W}) \\ \cdot \\ \cdot \\ \cdot \\ t_M(\hat{W}) \end{vmatrix} \quad (2.9)$$

Equation 2.8 is to be interpreted as a nonlinear system of  $M$  equations in  $N$  unknowns. Enough data  $\hat{\tau}$  must be given to allow smoothing of the experimental errors in the diffraction model. Consequently,  $M > N$  (or at the least  $M \geq N$ ) so the system in Eq. 2.8 is formally overdetermined.

### 3. NONLINEAR LEAST SQUARES: STRATEGY

We will attempt to "solve" the inverse problem by requiring that  $\hat{t}(\hat{W})$  matches  $\hat{t}$  in the sense that

$$E(\hat{W}) = \sum_{m=1}^M [\tau_m - t_m(\hat{W})]^2 \quad (3.1)$$

be minimized when considered as a function of  $\hat{W}$ . In other words, we approach the inversion as an unconstrained, nonlinear least squares problem. (The fact that  $w_{(N-1)/2} = 0$ , is not a real constraint since it amounts to translating the entire  $\hat{W}$  vector by a constant and it is shown in Appendix A that  $\hat{t}(\hat{W})$  is invariant under such a translation). Upon defining the vector

$$\hat{\phi}(\hat{W}) \equiv \hat{t}(\hat{W}) - \hat{t} \quad (3.2)$$

it is a simple matter to rewrite E as

$$E(\hat{W}) = \hat{\phi}^+(\hat{W})\hat{\phi}(\hat{W}) \quad (3.3)$$

Iterative methods will be employed in that we will replace the nonlinear problem of M equations in N unknowns by a sequence of linear least squares problems.

The Taylor series of the objective function  $E(\hat{W})$  can be used to approximate the minimum value of E from points  $\hat{W}$  near to the minimum  $\hat{W}_{\min}$  by setting  $\hat{W}_{\min} = \hat{W} + \Delta \hat{W}$ . Consequently

$$\begin{aligned} E(\hat{W}_{\min}) &\approx E(\hat{W}) + \sum_{n=1}^N \frac{\partial E(\hat{W})}{\partial w_n} \Delta w_n \\ &+ \sum_{j=1}^N \sum_{l=1}^N \frac{\partial^2 E(\hat{W})}{\partial w_j \partial w_l} \Delta w_l \end{aligned} \quad (3.4)$$

upon neglecting cubic and higher powers of  $\hat{\Delta \hat{W}}$ . The vector  $\hat{\Delta \hat{W}}$  is as yet unknown and our object is to determine this vector (of parameter corrections) which will approximate the minimum of E from the point  $\hat{W}$ .

As before, it will be convenient to work in matrix notation. Let us define  $\hat{g}$ , the Jacobian gradient vector, of E as

$$\hat{g} = \begin{vmatrix} \frac{\partial E}{\partial w_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial E}{\partial w_N} \end{vmatrix} \quad (3.5)$$

and  $\hat{H}$ , the Hessian matrix of E, as

$$\hat{H} = \begin{vmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial^2 E}{\partial w_1 \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_1 \partial w_N} \\ \frac{\partial^2 E}{\partial w_2 \partial w_1} & \frac{\partial^2 E}{\partial w_2^2} & \dots & \frac{\partial^2 E}{\partial w_2 \partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 E}{\partial w_N \partial w_1} & \frac{\partial^2 E}{\partial w_N \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_N^2} \end{vmatrix} \quad (3.6)$$

Note that  $\hat{H}$  is a symmetric NxN matrix. A further matrix which we will utilize is the Jacobian matrix

$$\hat{G} = \begin{vmatrix} \frac{\partial \phi_1}{\partial w_1} & \frac{\partial \phi_1}{\partial w_2} & \dots & \frac{\partial \phi_1}{\partial w_N} \\ \frac{\partial \phi_2}{\partial w_1} & \frac{\partial \phi_2}{\partial w_2} & \dots & \frac{\partial \phi_2}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \phi_M}{\partial w_1} & \frac{\partial \phi_M}{\partial w_2} & \dots & \frac{\partial \phi_M}{\partial w_N} \end{vmatrix} \quad (3.7)$$

This matrix is generally not square because  $M \geq N$ . Given these matrices, we can rewrite Eq. 3.4 in matrix form as

$$E(\hat{W}_{\min}) \approx E(\hat{W}) + \hat{g}^+ \Delta \hat{W} + \frac{1}{2} \Delta \hat{W}^+ \hat{H} \Delta \hat{W} \quad (3.8)$$

The elements of  $\hat{g}$  and  $\hat{H}$  can be expressed directly in terms of  $\phi_m$ . Now

$$\frac{\partial E}{\partial w_n} = 2 \sum_{m=1}^M \phi_m \frac{\partial \phi_m}{\partial w_n} \quad (3.9)$$

which can be cast into matrix form as

$$\hat{g} = 2\hat{G}^+ \hat{\phi} \quad (3.10)$$

The corresponding elements of the  $N \times N$  Hessian matrix of  $E$  are



$$\frac{\partial^2 E}{\partial w_j \partial w_l} = 2 \sum_{m=1}^M \frac{\partial \phi_m}{\partial w_j} \frac{\partial \phi_m}{\partial w_l} + 2 \sum_{m=1}^M \phi_m \frac{\partial^2 \phi_m}{\partial w_j \partial w_l} \quad (3.11)$$

with  $j, l = 1, 2, \dots, N$ . In some versions of least squares, the second term is neglected so that  $\hat{H}$  can be approximated by  $2\hat{G}^+\hat{G}$ . However, we will not make this approximation. See Appendix A for the explicit expressions in terms of our diffraction model, Eq. 2.4.

To determine that value of  $\Delta\hat{W}$  which makes  $E$  stationary, we equate to zero the gradient of  $E$  keeping  $\hat{g}$  and  $\hat{H}$  fixed. The result is

$$\hat{H}\Delta\hat{W} + \hat{g} = 0 \quad (3.12)$$

or

$$\hat{H}\Delta\hat{W} = -2\hat{G}^+\hat{\phi} \quad (3.13)$$

where  $\hat{H}$ ,  $\hat{G}$ , and  $\hat{\phi}$  are evaluated at  $\hat{W}$ . The solution  $\Delta\hat{W}$  of this system of linear equations gives the fundamental second order increment towards the minimum of  $E$ .

Equation 3.13 was established from a linearization of the basic system, Eq. 2.8. The nonlinear least square solution will (hopefully!) be reached after a sequence of iterations. After each iteration, Eq. 3.11, is used to obtain a new Hessian of partial derivatives and the process is repeated until the error metric  $E$  stabilizes (i.e., until no further diffraction model parameter improvements can be usefully made). Although this

procedure looks straightforward, it abounds with basic computational pitfalls which we discuss in the next section.

#### 4. NONLINEAR LEAST SQUARES: TACTICS

The previous section was devoted to strategy, the present section with the tactics of solving Eqs. 3.13 and 2.8. The basic difficulty in the calculations is the fact that the Hessian matrix  $\hat{H}$ , although  $N \times N$ , is at most of rank  $(N-1)$ . See Appendix B for details. Consequently, the formal inverse  $\hat{H}^{-1}$  does not exist and more sophisticated procedures must be employed in order to "invert" Eq. 3.13. Furthermore  $\hat{H}$  is ill conditioned (see Eq. 4.4).

We have chosen to employ the method of singular value decomposition (with an important modification) to evaluate the psuedo-inverse of  $\hat{H}$ . Singular value decomposition has among its several virtues the ability to determine the rank of  $\hat{H}$  during the computation. For those readers not familiar with singular value decomposition, we have outlined a version appropriate to the  $N \times N$  Hessian in Appendix C.

The solution to Eq. 3.13 is given by Eq. B.8

$$\Delta \hat{W} = - \sum_{\ell=1}^k \frac{\hat{u}_{\ell}^{+} \hat{G}^{+} \hat{\phi}}{\sigma_{\ell}} \hat{v}_{\ell}, \quad k \leq N \quad (4.1)$$

where  $\hat{u}_{\ell}$  and  $\hat{v}_{\ell}$  are the  $\ell^{\text{th}}$  column vectors of  $U$  and  $V$ , respectively. Equation 4.1 shows that the solution  $\Delta \hat{W}$  is a linear combination of  $k$  matrices,  $\hat{v}_{\ell}$ , each of rank one since  $\hat{u}_{\ell}^{+} \hat{G}^{+} \hat{\phi}$  is a scalar.

The ill-posed nature of the inversion is directly evident in Eq. 4.1. The smaller singular values entering into the denominator of the terms of the expansion tend to magnify greatly any error in the data vector  $\hat{\phi}$  resulting in a spurious solution. To correct this state of affairs, the expansion is terminated in a rational fashion before the contamination due to the numerically small singular values sets in.

Our choice to accomplish this is to use a filtering procedure. We rewrite Eq. 4.1 in the form

$$\Delta \hat{W} = - \sum_{\ell=1}^k f(\sigma_{\ell}) (\hat{u}_{\ell}^{+} \hat{G}^{+} \hat{\phi}) \hat{v}_{\ell} \quad (4.2)$$

where  $f(\sigma_{\ell})$  is a "filter function" depending on the singular values  $\sigma_{\ell}$ .  $f(\sigma_{\ell})$  is required to act like  $1/\sigma_{\ell}$  for large  $\sigma_{\ell}$ , approach 0 for very small  $\sigma_{\ell}$ ; and, finally, to decrease from  $1/\sigma_{\ell}$  to zero smoothly in the intermediate range. A useful candidate, evidently first used by Crone [13], is

$$f(\sigma_{\ell}) = \frac{\sigma_{\ell}^L}{\sigma_{\ell}^{L+1} + q^{L+1}} \quad (4.3)$$

where  $L$  is some non-negative integer and  $q$  is a non-negative real parameter.

Another way to view the desirability of using a filter function is to note that the components of  $\hat{\Delta W}$  that lie in the direction corresponding to large singular values are the important ones insofar as reducing  $E$ . It is the remaining components of  $\hat{\Delta W}$  that cause numerical instability because they lie in directions that allow very large changes with little effect on the actual approximation. The filter function given in Eq. 4.3 attenuates these parasitic components of  $\hat{\Delta W}$ . Thus,  $q$  plays the role of a variable metric and is to be chosen reasonably large,  $q=0(.1)$ , when far from the least squares solution and decreased as the iterations are sequenced.

Given an initial estimate (really a guess!) of  $\hat{W}$ , say  $\hat{W}^{(1)}$ ; we first perform a singular value decomposition of the Hessian matrix  $\hat{H}$  evaluated at  $\hat{W}^{(1)}$ . We solve for the search direction  $\hat{\Delta W}$  using Eq. 3.13 employing a predetermined filter function and  $q$  and search along this direction a distance  $s_k$  for the minimum  $\hat{W}^{(2)}$ . Unlike the usual steepest descent method which forces one to search for a minimum in the direction of the negative gradient of  $E$ , the present method modifies this direction by the pseudoinverse of the Hessian, which contains gradient and slope information. The new estimate  $\hat{W}^{(2)}$  is then iterated to obtain  $\hat{W}^{(3)}$  decreasing  $q$  as  $\|\hat{\Delta W}\|$  decreases.

Due to changes in  $\hat{H}$ , the filter must be modified at each

stage in actual calculations. The Hessian  $\hat{H}$  is almost always ill conditioned in the sense that

$$\sigma_{\max} / \sigma_{\min} \gg 1 \quad (4.4)$$

where  $\sigma_{\max}$  and  $\sigma_{\min}$  are the largest and smallest *positive* singular values.

The first consideration is scaling,  $\sigma_{\max}$  varies with  $M$ ,  $N$  and  $k$ . Since ill conditioning is defined relative to  $\sigma_{\max}$ , the filter function  $f(\sigma_\ell)$  defined in Eq. 4.3 must also be defined relative to  $\sigma_{\max}$  by appropriate scaling. We set  $L = 2$  in Eq. 4.3 in all the calculations reported in this paper. Furthermore,  $f(\sigma_\ell)$  is scaled thusly

$$f(\sigma_\ell) = \frac{1}{\sigma_{\max}} \frac{p_\ell^2}{p_\ell^3 + q_k^3} \quad (4.5)$$

where  $p_\ell = (\sigma_\ell / \sigma_{\max})$ . The scaled filter is made to depend upon the  $k$ th iteration by choosing  $q$  in Eq. 4.3 to depend on  $k$  (i.e.,  $q \rightarrow q_k$ ).

A reasonable definition of  $q_k$  must depend upon the following two requirements:

1. As  $\hat{W}^{(k)}$  converges to the minimum, we want the filtered inverse Hessian to converge to the *true* inverse Hessian (i.e.,  $f(\sigma_\ell) \rightarrow \sigma_\ell^{-1}$ ). We monitor such convergence by noting the size

of the relative change  $\gamma_k$  in  $\hat{W}^{(k)}$  defined by

$$\gamma_k = \frac{\|s_k \Delta \hat{W}^{(k)}\|}{\|\hat{W}^{(k)}\|} \quad (4.6)$$

From theory [14,15], the convergence near the minimum should be quadratic, so the filter convergence is made superlinear to take some advantage of this fact.

2. As shown in Appendix D, a minimizer  $\hat{W}^{(0)}$  exists in the region  $\{\hat{W}: -\pi \leq w_n \leq \pi\}$ . Consequently it is desirable to keep  $\Delta \hat{W}^{(k)}$  of this order of magnitude to ensure iterates stay in this region. So if  $\gamma_k$  is large, we choose the  $q_k$  to pass only the larger singular values.

With these considerations in mind,  $q_k$  is chosen to be

$$q_k \equiv \min \left\{ .1, 30(\gamma_k)^{1.5} \right\} \quad (4.7)$$

These parameters were found by trial and error.

A deeper understanding of the effects and advantages of filtering comes from the observation that filtering is equivalent to Tikhonov regularization [11,12]. If no filter is used then by statements A and B of Appendix C,  $\Delta \hat{W}^{(k)}$  will minimize  $\|\hat{H}^{(k)} \Delta \hat{W}^{(k)} + \hat{g}^{(k)}\|$ . If a filter is used, then there exists a

scalar  $\lambda_k$  and matrix  $\hat{A}^{(k)}$  such that  $\Delta \hat{W}^{(k)}$  will minimize

$$\| \hat{H}^{(k)} \Delta \hat{W}^{(k)} + \hat{g}^{(k)} \| + \lambda_k \| \hat{A}^{(k)} \Delta \hat{W}^{(k)} \| \quad (4.8)$$

i.e.,  $\Delta \hat{W}^{(k)}$  is a regularized least squares solution. According to the Tikhonov approach the regularized solution that minimizes Eq. 4.8 will be "smoother" than the least squares solution.

In the method used here although each linear problem is regularized by filtering, as  $k$  increases  $\lambda_k \rightarrow 0$  so the nonlinear problem is not regularized, i.e., we are minimizing iteratively the nonlinear function  $E(\hat{W})$ , *not* the regularized function  $E(\hat{W}) + \lambda \| \hat{W} \|^2$ . Viewing the filter as a regularizer shows us clearly that filtering places emphasis on those directions in which  $E(\hat{W})$  is fastest in decreasing, and neglects directions in which  $E(\hat{W})$  is slowly decreasing, thereby trading these off for a  $\Delta \hat{W}$  of small norm. Thus  $\hat{W}^{(k)}$  stays as close as possible to  $\hat{W}^{k-1}$ , a good policy if  $\hat{W}^{(1)}$  is a good initial guess.

Newton's method performs well as long as care is taken in inverting the Hessian and a "good" initial guess for  $\hat{W}^{(0)}$  is available. There are bounds on an initial guess, expressed in terms of the higher derivatives of  $E(\hat{W})$ , which guarantee the convergence of Newton's method from that guess; this is the essence of the Newton-Kantorovich theorem [15,16]. Such bounds are



generally very conservative and also hard to evaluate. However due to the simple analytical form of  $E(\hat{W})$  the calculations can be done although we have as yet not performed this task. There is an algorithm due to Kung [17] which makes use of these bounds to generate a succession of Newton iterates that will converge from any initial guess.

There are available other methods which produce a sequence that decreases  $E(\hat{W})$  until a sufficiently good starting point for Newton's method is reached (e.g., the simplex method of Nelder and Mead [18] as described in Daniels [19]). However the difficulty is to determine when such a point is reached without calculating the bounds mentioned above.

To circumvent this aspect of the problem and avoid using another minimizing algorithm, we applied the following procedure:

1. Run Newton's method for  $N$  very small from an arbitrary initial guess to achieve a minimum  $\hat{W}_N^{(0)}$ .
2. Increase  $N$  to  $N'$ , produce  $\hat{W}_{N'}^{(1)}$  by interpolation over  $\hat{W}_N^{(0)}$ .
3. Run Newton's method with  $\hat{W}_{N'}^{(1)}$  as an initial guess until it converges on  $\hat{W}_{N'}^{(0)}$ . Return to 2.

The algorithm is repeated until a sufficiently large  $N$  is reached. In practice it was found that  $N = 5$  was a reasonable starting  $N$  value. However  $N'$  could not be increased markedly over  $N$ . Limits on available computation time did not allow us to experiment to determine optimal procedures.

The final "tactical" problem to be discussed is the stopping rule (i.e., the criterion used to decide when  $\hat{W}^{(k)}$  is sufficiently close to  $\hat{W}^{(0)}$  so that the iterative algorithm can be halted). The stopping rule was made independent of  $E(\hat{W})$  for two reasons. The first is that the shape and magnitude of the surface defined by  $E(\hat{W})$  depends on  $N, M$  and the measurement noise in  $\hat{r}$ , so that criteria using  $E(\hat{W})$  would be too problem dependent to allow comparison of results.

The second reason is that  $E(\hat{W})$  can be insensitive to large changes in  $\hat{W}$  (this is indicated by the ill conditioning of  $\hat{H}$ ). Thus a stopping rule based on changes in  $E(\hat{W})$  only, can halt the iterative algorithm far from a minimum.

For these reasons, a relative error based on  $\hat{W}^{(k)}$  was first considered namely: halt iteration algorithm when

$$\gamma_k < \epsilon \quad (4.9)$$

where  $\gamma_k$  is given by Eq. 4.6 and  $\epsilon$  is a specified constant. Trial calculations showed that the iterative algorithm often took a small step followed by a large step (typical behavior of a minimization algorithm descending a long, curving valley with steep walls, e.g., Rosenbrook function [20]). Thus to avoid stopping after a small step while still far from the minimum, the stopping rule finally adopted was: halt iteration algorithm when

$$\gamma_{k+1} + \gamma_k < \epsilon \quad (4.10)$$

After some numerical experimentation,  $\epsilon = .005$  was chosen.

## 5. SOME NUMERICAL RESULTS

To test the numerical workings of the entire algorithm, we considered a known wavefront aberration function  $W(p)$  and from it calculated the diffraction model point spread function  $t(v_m)$  using Eq. 2.4. Noise was introduced into the measured point spread function  $\hat{t}$  in a multiplicative fashion

$$\hat{t}_{\text{noisy}} = (1 + \delta\mu)\hat{t} \quad (5.1)$$

where  $\delta$  is a positive constant less than unity and  $\mu$  is a random variable uniformly distributed over  $(-1, +1)$

$$\begin{aligned} f(\mu) &= \frac{1}{2} & |\mu| < 1 \\ &= 0 & |\mu| > 1 \end{aligned} \quad (5.2)$$

Values of  $\delta$  used in the present calculations are  $\delta = 0.025$  and  $\delta = .05$  described loosely as 5% and 10% noise.

The sampled values  $v_m$  were taken to be  $v_m = m\pi/2$  in accordance with the sampling expansion [21] appropriate to slit aperture. Furthermore, the number  $M$  of sampled values was taken to be odd in order to take the maximum value of  $\hat{t}$  as  $m = 0$ . All calculations reported are for  $N = 21$  and  $M = 21$  or  $31$ .

The wavefront aberration function was taken to consist of coma and spherical aberration

$$\begin{aligned}
w(p) &\equiv \frac{2\pi}{\lambda} W(p) \\
&= 2\pi[W_3 S_3(p) + W_4 S_4(p)] \\
&= 2\pi[W_3(p^3 - \frac{3}{5} p) + W_4(p^4 - \frac{6}{7} p^2)] \quad (5.3)
\end{aligned}$$

where  $W_3$  and  $W_4$  are measured in wavelength units (i.e.,  $W_3/\lambda$ ,  $W_4/\lambda$  are dimensionless).  $S_3(p)$  for coma and  $S_4(p)$  for spherical aberration are the slit aperture versions [22] of the Zernike polynomials. The numerical calculations were carried out for

$$W_3 = W_4 = \frac{3}{8} \lambda \quad (5.4)$$

The true wavefront aberration function is shown as a solid line in the succeeding figures.

Our problem is to determine the wavefront aberration function from the noisy sampled point spread function, Eq. 5.1, and compare it with the true wavefront, Eq. 5.3.

In the first set of calculations, the noise level was set at 5% and the number of sampling points  $M$  was taken to be 31. Following the procedure described in the previous section, an initial guess was iterated until a satisfactory  $E$  was achieved as per the stopping algorithm discussed in the previous section with  $\epsilon = .005$ . Approximately twenty iteratives were performed to achieve these levels of  $E$ . The results of three typical sample

realization reconstructions of  $w(p)$  are shown in Figs. 1-3 along with the corresponding values of  $E$ . The reconstructed values are in excellent agreement with the true values and do not require any detailed comment.

In order to test the stability of the algorithm with respect to the number of sampling points, we next set  $M=21$  (so that the number of sampling points equals the number of reconstructed wavefront points) and kept the same noise level of 5%. Two sample realization reconstructions are displayed in Figs. 4,5. Overall these results are of about the same accuracy as those with more sampling points.

Finally, we ran calculations for 10% noise with 31 sampling points. Two reconstructions are shown in Figs. 6 and 7. The results in Fig. 6 are extremely good, even those in Fig. 7 are respectable.

It is of some interest to list the final iterate singular values corresponding to Figs. 1-3, see Table 1. As discussed in Appendix B, the Hessian is singular which is reflected in the fact that  $\sigma_{21} = 0$ . The first few ordered singular values are roughly equal for the three cases in question; however, the higher order  $\sigma_k$  are very small and highly irregular in their behavior as can be seen by comparison of the first and third columns. The

other cases behave in much the same manner, always with  $\sigma_{21} \equiv 0$ .

Calculations were also performed on two special forms of  $W(p)$ :

- a. even parity,  $W(p) = W(-p)$ , (e.g., spherical aberration)
- b. odd parity,  $W(p) = -W(-p)$ , (e.g., coma).

Observed behavior led to the establishment of the following results (the proofs are omitted for brevity).

1. If  $\hat{W}^{(k)}$  is odd, then all subsequent  $\hat{W}^{(\ell)}$  are odd for  $\ell > k$ .
2. If  $\hat{W}^{(k)}$  is even and  $\hat{t}$  is even, then all subsequent  $\hat{W}^{(\ell)}$  are odd for  $\ell > k$ .
3. If  $\hat{W}^{(0)}$  minimize  $E(\hat{W})$  and  $\hat{W}^{(0)}$  is even, then  $-\hat{W}^{(0)}$  also minimizes  $E(\hat{W})$ .

An odd or even  $\hat{W}^{(k)}$  does not imply that the Hessian  $\hat{H}[\hat{W}^{(k)}]$  is degenerate, rather that the spaces of odd vectors and even vectors are eigenspaces of  $\hat{H}$ . If, for example, at any stage  $\hat{W}^{(k)}$  is odd then condition 2 implies that Newton's method is henceforth restricted to a  $(N-1)/2$  dimensional subspace of possible solutions which may not contain the true minimum of  $E(\hat{W})$ , even though the range of  $\hat{H}(\hat{W})$  may be larger than this subspace. Condition 3 raises the point that although singular value decomposition

and filtering ensure that  $\Delta \hat{W}^{(k)}$  is uniquely defined for each subproblem, the problem as a whole can have several solutions, each of which is a potential point of convergence for Newton's method.

Rapid convergence from any initial state was observed for  $\hat{t}$ 's calculated from even or odd  $W(p)$ ; especially so if the initial guess has the same parity. In some of these cases the rank of  $\hat{H}$  reduced to  $(N-1)/2$ .



## APPENDIX A

The matrices  $\hat{g}$ ,  $\hat{G}$  and  $\hat{H}$  defined in Section 3 are given in terms of derivatives of  $E$  via Eqs. 3.9 and 3.11. The derivatives of  $E$ , in turn, require a knowledge of  $\hat{\phi}(\hat{W})$  and its first two derivatives.

The explicit expression for  $\hat{\phi}(\hat{W})$ , as defined in Eq. 3.2, for our model is

$$\begin{aligned}\phi_m(\hat{W}) &= \left| \frac{1}{2} \sum_{n=1}^N \alpha_n e^{iv_m p_n} e^{iw_n} \right|^2 - \tau_m \\ &= \frac{1}{4} \left\{ \left[ \sum_{n=1}^N \alpha_n \cos(v_m p_n + w_n) \right]^2 \right. \\ &\quad \left. + \left[ \sum_{n=1}^N \alpha_n \sin(v_m p_n + w_n) \right]^2 \right\} - \tau_m \quad (A.1)\end{aligned}$$

The first derivative becomes

$$\begin{aligned}\frac{\partial \phi_m}{\partial w_k} &= \frac{1}{2} \alpha_k \cos(v_m p_k + w_k) \sum_{n=1}^N \alpha_n \sin(v_m p_n + w_n) \\ &\quad - \frac{1}{2} \alpha_k \sin(v_m p_k + w_k) \sum_{n=1}^N \alpha_n \cos(v_m p_n + w_n) \quad (A.2)\end{aligned}$$

where  $k=1, \dots, N$ . The second partial derivatives of  $\phi_m$  are

$$\begin{aligned}
\frac{\partial^2 \phi_m}{\partial w_k \partial w_l} &= \frac{1}{2} \alpha_k \alpha_l \cos[v_m(p_k - p_l) + (w_k - w_l)] \quad , \quad k \neq l \\
&= -\frac{1}{2} \sum_{n=1}^{N'} \alpha_k \alpha_n \cos[v_m(p_k - p_n) + (w_k - w_n)] \quad , \quad k=l
\end{aligned}
\tag{A.3}$$

where the prime on the summation sign implies that the term  $n=k$  is to be omitted.

## APPENDIX B

A major factor in the difficulty of numerically finding a minimizer  $\hat{W}^{(0)}$  (say) of  $E(\hat{W})$  is that many such minimizers exist. Consequently each is a potential point of attraction for the algorithm. In this appendix the existence of a one-dimensional subspace of minimizers is demonstrated; furthermore this ensures that the rank of the Hessian  $\hat{H}(\hat{W})$  is always less than or equal to  $(N-1)$ .

It is convenient to define the quantities

$$\begin{aligned} c_{m,n} &\equiv \alpha_n \cos(v_m p_n + w_n) \\ s_{m,n} &\equiv \alpha_n \sin(v_m p_n + w_n) \\ C_m &\equiv \sum_{n=1}^N c_{m,n} \\ S_m &\equiv \sum_{n=1}^N s_{m,n} \end{aligned} \tag{B.1}$$

From these definitions and the relevant expressions in Appendix A, it follows that

$$\phi_m(\hat{W}) = \frac{1}{4} (C_m^2 + S_m^2) - \tau_m \tag{B.2}$$

$$\frac{\partial \phi_m(\hat{W})}{\partial W_n} = \frac{1}{2} (S_m c_{m,n} - C_m s_{m,n}) \quad (B.3)$$

The Hessian  $\hat{H}$  can be written as

$$H_{np} = 2 \sum_{m=1}^M \left[ A_{np}^{(m)} + \phi_m B_{np}^{(m)} \right] \quad (B.4)$$

where

$$\begin{aligned} A_{np}^{(m)} &\equiv \frac{\partial \phi_m}{\partial W_n} \frac{\partial \phi_m}{\partial W_p} \\ &= \frac{1}{4} (S_m^2 c_{m,n} - S_m C_m s_{m,n}) c_{m,p} \\ &\quad + \frac{1}{4} (C_m^2 s_{m,n} - S_m C_m c_{m,n}) s_{m,p} \end{aligned} \quad (B.5)$$

and

$$\begin{aligned} B_{np}^{(m)} &\equiv \frac{\partial^2 \phi_m(\hat{W})}{\partial W_n \partial W_p} \\ &= \frac{1}{2} (c_{m,p} c_{m,n} + s_{m,p} s_{m,n}) \quad , \quad n \neq p \\ &= \frac{1}{2} (c_{m,n}^2 + s_{m,n}^2 - S_m s_{m,n} - C_m c_{m,n}) \quad , \quad n=p \end{aligned} \quad (B.6)$$

Let  $\hat{e}$  be a vector whose entries are all unity. We now show that if  $\hat{W}^{(0)}$  minimizes  $E(\hat{W})$ , then so will the vector  $\hat{W}^{(0)} + c\hat{e}$  for any real  $c$ . The proof will follow from the result that

$$\hat{\phi}(\hat{W}) = \hat{\phi}(\hat{W} + c\hat{e}) \quad \forall \hat{W} \quad (B.7)$$

Furthermore this result causes the Hessian to have rank at most (N-1) by proving that  $\hat{e}$  is in the null space of  $\hat{H}(\hat{W})$ , i.e.,

$$\hat{H}(\hat{W})\hat{e} = \hat{0} \quad , \quad \forall \hat{W} \quad (B.8)$$

This is equivalent to showing that the determinant of  $\hat{H}(\hat{W})$  vanishes. The formal proof follows by virtue of two lemmas.

*Lemma 1*

$$\hat{\phi}(\hat{W}) = \hat{\phi}(\hat{W} + c\hat{e}) \quad (B.9)$$

This follows from

$$\begin{aligned} \phi_m(\hat{W} + c\hat{e}) &= \left| \frac{1}{2} \sum_{n=1}^N e^{iv_m p_n} e^{iw_n} e^{ic} \right|^2 \\ &= \left| e^{ic} \right|^2 \phi_m(\hat{W}) \\ &= \phi_m(\hat{W}) \end{aligned} \quad (B.10)$$

*Lemma 2*

$$\hat{H}(\hat{W})\hat{e} = \hat{0} \quad (B.11)$$

For the proof, it suffices to show that

$$\hat{A}^{(m)}_e = \hat{0} \quad (B.12)$$

and

$$\hat{B}^{(m)}_e = \hat{0} \quad (B.13)$$

Now

$$\begin{aligned} (\hat{A}^{(m)}_e)_n &= \sum_{p=1}^N A_{np}^{(m)} \\ &= \sum_{p=1}^N \frac{1}{4} (S_m^2 c_{m,n} - S_m C_m s_{m,n}) c_{m,p} \\ &\quad + \sum_{p=1}^N \frac{1}{4} (C_m^2 s_{m,n} - C_m S_m c_{m,n}) s_{m,p} \\ &= \frac{1}{4} (S_m^2 c_{m,n} - S_m C_m s_{m,n}) C_m \\ &\quad + \frac{1}{4} (C_m^2 s_{m,n} - C_m S_m c_{m,n}) S_m \\ &= 0 \end{aligned} \quad (B.14)$$

Also

$$\begin{aligned}
 (\hat{B}^{(m)} \hat{e})_n &= \sum_{p=1}^N B_{np}^{(m)} \\
 &= \sum_{p=1}^N \frac{1}{2} (c_{m,p} c_{m,n} + s_{m,p} s_{m,n}) \\
 &\quad - \frac{1}{2} (s_{m,p} s_{m,n} + c_{m,p} c_{m,n}) \\
 &= \frac{1}{2} (c_{m,p} c_{m,n} + s_{m,p} s_{m,n}) \\
 &\quad - \frac{1}{2} (s_{m,p} s_{m,n} + c_{m,p} c_{m,n}) \\
 &= 0
 \end{aligned} \tag{B.15}$$

In the program a particular  $\hat{W}^{(0)}$  of this subspace is chosen due to the constraint  $w(0) = 0$  or in discrete form  $w_{(N-1)/2} = 0$ . Since given any minimum  $\hat{W}^{(0)}$ , a  $\hat{W}^{(1)}$  which satisfies the above constraint can be constructed by choosing a particular constant  $c$  and letting

$$\hat{W}^{(1)} = \hat{W}^{(0)} + c \hat{e} \tag{B.16}$$

then the constraint  $w(0) = 0$  can never be a binding constraint. However, choice of a particular  $\hat{W}^{(0)}$  does not remove the difficulties caused by the existence of an infinite set of such  $\hat{W}^{(0)}$ .



## APPENDIX C

The Hessian matrix  $\hat{H}$ , Eq. 3.6, can be expressed as the product of three matrices (singular value decomposition of  $\hat{H}$ )

$$\hat{H} = \hat{U} \hat{S} \hat{V}^+ \quad (C.1)$$

where  $\hat{U}$  and  $\hat{V}$  are  $N \times N$  orthogonal matrices (ie,  $\hat{U} \hat{U}^+ = \hat{U}^+ \hat{U} = \hat{I}$ , the same for  $\hat{V}$ ) and  $\hat{S}$  is an  $N \times N$  diagonal matrix.

$$\hat{S} = \begin{vmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_N \end{vmatrix} \quad (C.2)$$

The  $\sigma$ 's are termed the *singular values* and are the eigenvalues of

$$\hat{H}^+ \hat{H} \hat{S}_l = \sigma_l \hat{S}_l \quad l = 1, 2, \dots, N \quad (C.3)$$

This is the *mathematical* definition of the singular values, but they are *calculated* by an entirely different procedure which guarantees their numerical stability. The  $\sigma$ 's can be numerically ordered

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_N \geq 0 \quad (C.4)$$

If  $\hat{H}$  is of rank  $k$ , where  $k < N$ , then the last  $N-k$  of the  $\sigma$ 's are zero.

The solution to the minimal least squares problem posed in Eq. 3.7 can be cast directly into a form involving the singular values and their corresponding singular vectors. Substitution of Eq. 3.1 into Eq. 3.13 yields after some matrix manipulations

$$\Delta \hat{W} = 2(\hat{V} \hat{S}^{\oplus} \hat{U}^+) \hat{G}^+ \hat{\phi} = -2 \hat{H}^{\oplus} \hat{G}^+ \hat{\phi} \quad (C.5)$$

The matrix  $\hat{H}^{\oplus} \equiv \hat{V} \hat{S}^{\oplus} \hat{U}^+$  is termed the Moore-Penrose pseudoinverse of  $\hat{H}$ . Here

$$\hat{S}^{\oplus} \equiv \begin{vmatrix} \sigma_1^+ & & & \\ & \sigma_2^+ & & \\ & & \ddots & \\ & & & \sigma_N^+ \end{vmatrix} \quad (C.6)$$

with

$$\begin{aligned} \sigma_n^+ &= \frac{1}{\sigma_n} & \text{if } \sigma_n > 0 \\ &= 0 & \text{if } \sigma_n = 0 \end{aligned} \quad (C.7)$$

It is not our intent to give a full discussion of the Moore-Penrose pseudoinverse for details are available in the literature [23-24]. Suffice it to say that it produces a  $\Delta \hat{W}$  which satisfies the two minimum conditions with respect to our *linear* problem:

A) It achieves the unique minimum of

$$\| \hat{H} \Delta \hat{W} + 2 \hat{G}^+ \hat{\phi} \|.$$

B) If there are any other  $\Delta \hat{W}$  which satisfy Eq. 3.13, then

Eq. C.5 is characterized among them by having the smallest norm; in other words, Eq. C.5 minimizes  $\| \Delta \hat{W} \|$  among the solutions.

The solution given by Eq. C.5 becomes somewhat more transparent if the right hand side of Eq. B.5 is written out more explicitly

$$\Delta \hat{W} = - \sum_{l=1}^k \frac{\hat{u}_l^+ \hat{G}^+ \hat{\phi}}{\sigma_l} \hat{v}_l, \quad k \leq N \quad (C.8)$$

where  $\hat{u}_l$  and  $\hat{v}_l$  are the  $l$ -th column vectors of  $\hat{U}$  and  $\hat{V}$  respectively.

## APPENDIX D

In this appendix we establish bounds on the minimum  $\hat{W}^{(0)}$ . We first prove that  $E(\hat{W})$  is periodic in each coordinate with period  $2\pi$ , i.e.,

$$E(\hat{W}) = E(\hat{W} + 2\pi \hat{e}^l) \quad (D.1)$$

where  $\hat{e}^l$  is the  $l$ th unit vector. For a proof it suffices to show the result for the components of  $\hat{\phi}(\hat{W})$ , thus

$$\begin{aligned} \phi_m(\hat{W} + 2\pi \hat{e}^l) &= \left| \sum_{n=1}^N e^{i v_m p_n} e^{i(w_n + 2\pi e_n^l)} \right|^2 \\ &= \left| \sum_{n=1}^N e^{i v_m p_n} e^{i w_n} \right|^2 \\ &= \phi_m(\hat{W}) \end{aligned} \quad (D.2)$$

This property is dependent on the particular discretization chosen. It establishes the desirable result that there exists a minimum  $\hat{W}^{(0)}$  of  $E(\hat{W})$  in the region

$$S = \{ \hat{W} \mid -\pi \leq w_n \leq \pi \} \quad (D.3)$$

Since  $E(\hat{W})$  has period  $2\pi$ , we have

$$\min_{\hat{W} \in \mathbb{R}^n} E(\hat{W}) = \min_{\hat{W} \in S} E(\hat{W}) \quad (D.4)$$

$E(\hat{W})$  is continuous and  $S$  is a closed set, consequently  $E(\hat{W})$  attains its minimum on  $S$  at some point  $\hat{W}^{(0)}$  in  $S$ . By periodicity,  $\hat{W}^{(0)}$  is also a global minimum.

Periodicity provides bounds for the minimum  $\hat{W}^{(0)}$  which are helpful in searches for  $\hat{W}^{(0)}$ , but at the same time indicates  $\hat{E}(\hat{W})$  is a complicated surface with many maxima, minima, and saddle points, obviously a surface on which most minimization algorithms will have difficulty!

## REFERENCES

1. R.A. Gonsalves, Phase retrieval from modulus data, *J. Opt. Soc. Am.*, **66**, 1977, 961.
2. W.H. Southwell, Wavefront analyzer using a maximum likelihood algorithm, *J. Opt. Soc. Am.*, **67**, 1977, 396.
3. A. Walther, The question of phase retrieval in optics. *Opt. Acta*, **10**, 1963, 41.
4. M.L. Goldberger, H.W. Lewis and K.M. Watson, Use of intensity correlations to determine the phase of a scattering amplitude. *Phys. Rev.*, **132**, 1963, 2764.
5. H. Nussenzveig, Phase problem in coherence theory. *J. Math. Phys.*, **8**, 1967, 561.
6. R.W. Gerchberg and W.O. Saxton, A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, **40**, 1972, 237.
7. J.R. Fienup, Reconstruction of an object from the modulus of its Fourier transform. *Opt. Lett.*, **3**, 1978, 27.
8. R.E. Burge, M.A. Fiddy, M. Nieto-Vesperinas, and M.W.L. Wheeler, The phase problem in scattering theory: the zeros of entire functions and their significance. *Proc. Roy. Soc. Lond.*, **A360**, 1978, 25.

9. W.O. Saxton, *Computer Techniques for Image Processing in Electron Microscopy* (Academic Press, New York, 1978).
10. H.P. Baltes, ed., *Inverse Source Problems in Optics* (Springer-Verlag, Berlin, 1978).
11. M. Lavrentiev, *On Some Improperly Posed Problems of Mathematical Physics* (Springer-Verlag, Berlin, 1963).
12. A.N. Tikhonov and V.Y. Arsenin, *Solutions of Ill-Posed Problems* (Halsted Press, New York, 1977).
13. L. Crone, The singular value decomposition of matrices and cheap numerical filtering of systems of linear equations. *J. Franklin Inst.*, 294, 1972, 133.
14. G. Dahlquist and A. Björck, *Numerical Methods* (Prentice-Hall, Englewood Cliffs, NJ, 1974).
15. J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Non-Linear Equations in Several Variables* (Academic Press, New York, 1970).
16. J.M. Ortega, The Newton-Kantorovich theorem. *Amer. Math. Monthly*, 75, 1968, 658.
17. H.T. Kung, The complexity of obtaining starting points for solving operator equations by Newton's method, in *Analytical Computational Complexity*, edited by J.F. Traub (Academic Press, New York, 1976).

18. J. Nelder and R. Mead, A simplex method for function minimization. *Compt. J.*, 7, 1965, 308.
19. R.W. Daniels, *An Introduction to Numerical Methods and Optimization Techniques* (North-Holland, New York, 1978).
20. P.R. Aday and M.A.H. Dempster, *Introduction to Optimization Methods* (Halsted Press; New York, 1977), p. 116.
21. R. Barakat, Application of the sampling theorem to optical diffraction theory. *J. Opt. Soc. Am.*, 54, 1964, 920.
22. R. Barakat and L. Reisberg, Diffraction theory of the aberrations of a slit aperture. *J. Opt. Soc. Am.*, 55, 1965, 878.
23. C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems* (Prentice-Hall, Englewood Cliffs, NJ, 1974) Chap. 7.
24. G. Golub and C. Reinsch, Singular value decomposition and least squares solutions. *Numer. Math.*, 14, 1970, 403. The numerical calculations in the present paper utilized the Golub-Reinsch algorithm for determining the singular values.

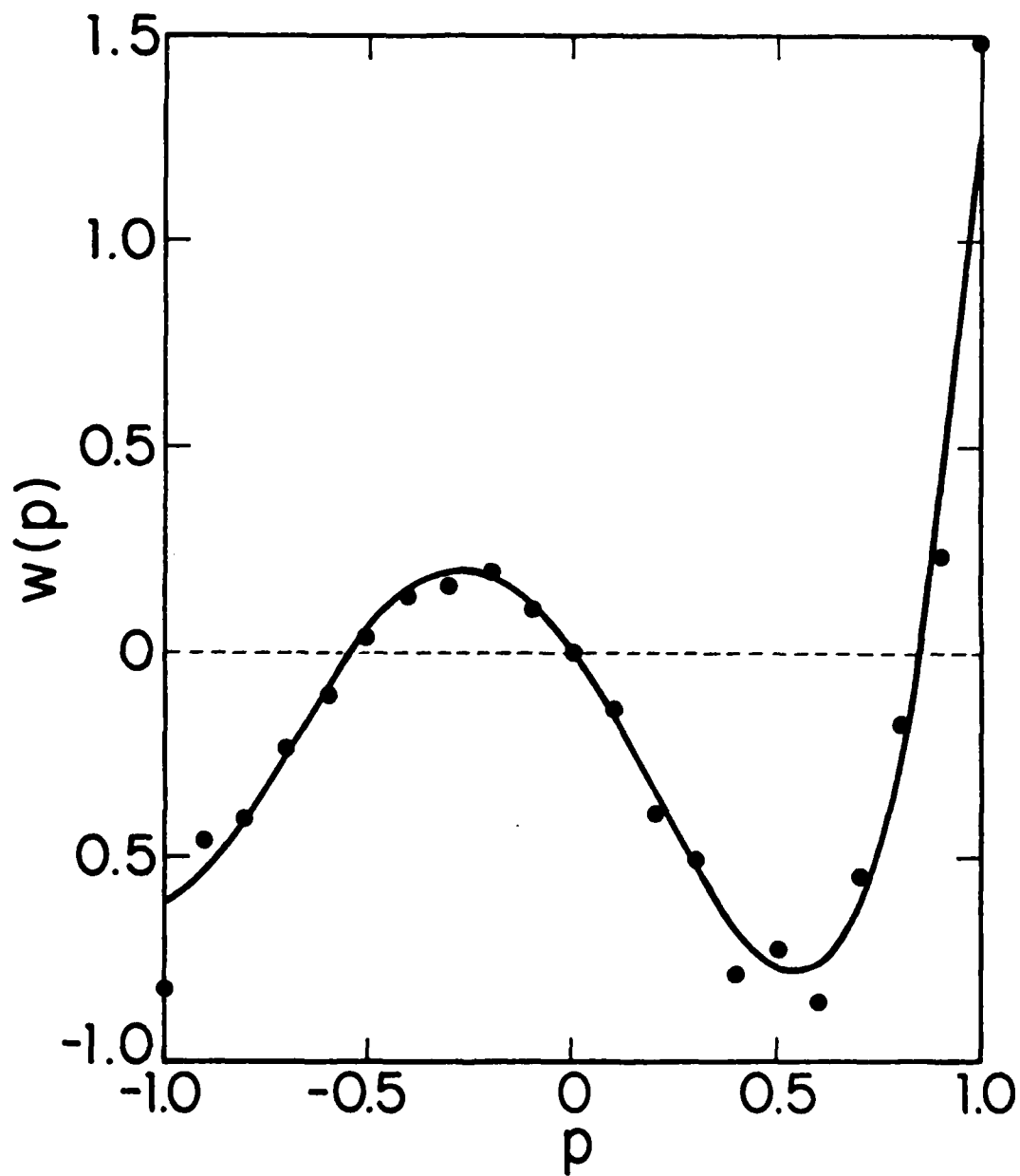


TABLE 1.: FINAL ITERATE SINGULAR VALUES  $\sigma_\ell$  CORRESPONDING  
TO FIGS. 1-3 RESPECTIVELY

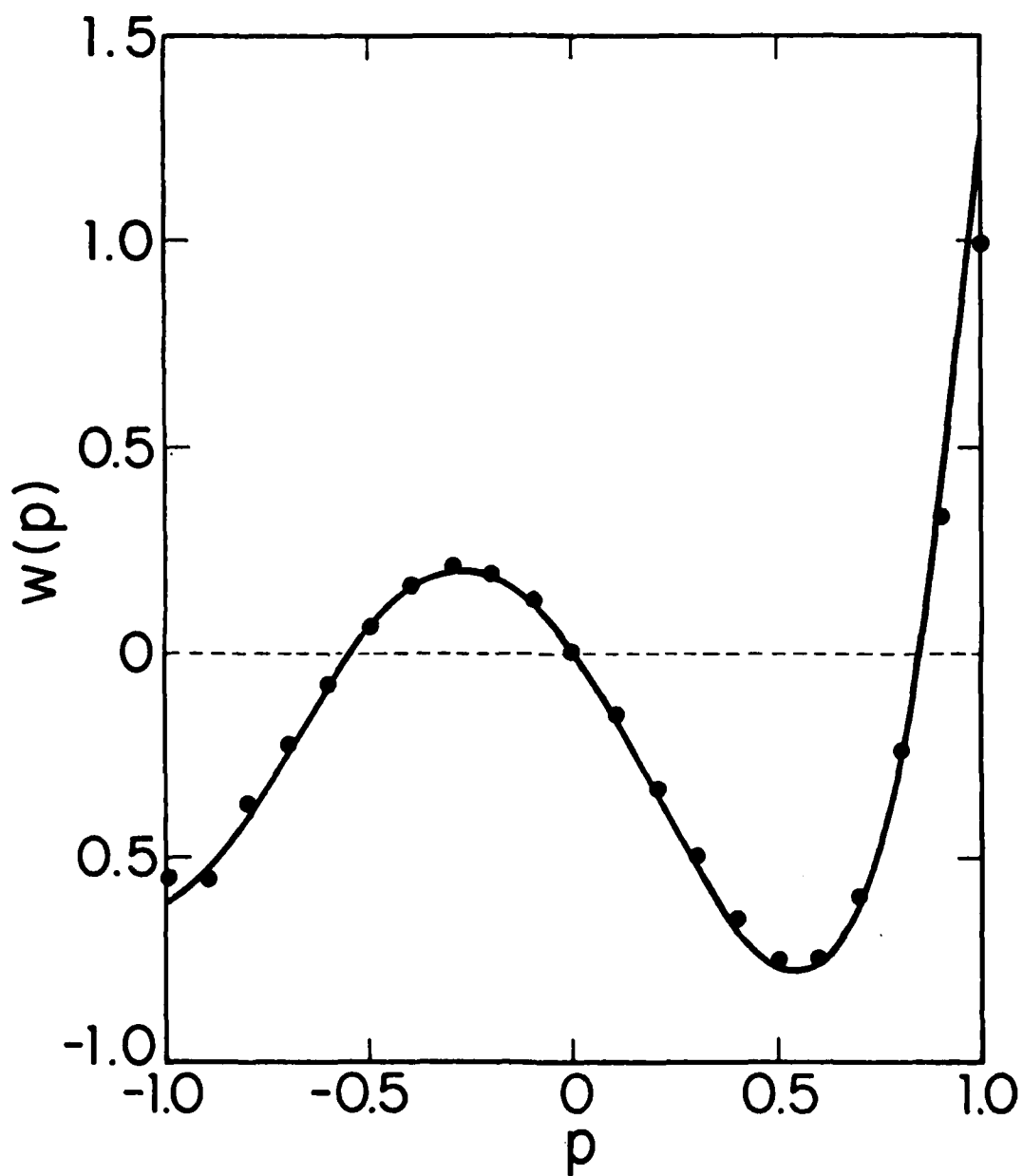
$\ell$	$\sigma_\ell$	$\sigma_\ell$	$\sigma_\ell$
1	.1299778	.1293188	.1308360
2	.0654607	.0663010	.0651481
3	.0147847	.0148309	.0151498
4	.0069802	.0059728	.0061527
5	.0044707	.0040800	.0040639
6	.0029514	.0025031	.0030560
7	.0022897	.0014872	.0021063
8	.0019339	.0014096	.0016379
9	.0014224	.0010693	.0014751
10	.0010348	.0008257	.0013507
11	.0009966	.0007715	.0012386
12	.0005798	.0005468	.0011485
13	.0003987	.0005114	.0010666
14	.0003521	.0004359	.0010383
15	.0002837	.0003983	.0009781
16	.0002275	.0003603	.0009411
17	.0001589	.0003352	.0008945
18	.0000700	.0003189	.0008881
19	.0000346	.0002814	.0005131
20	.0000170	.0002568	.0004882
21	0	0	0

## FIGURE LEGENDS

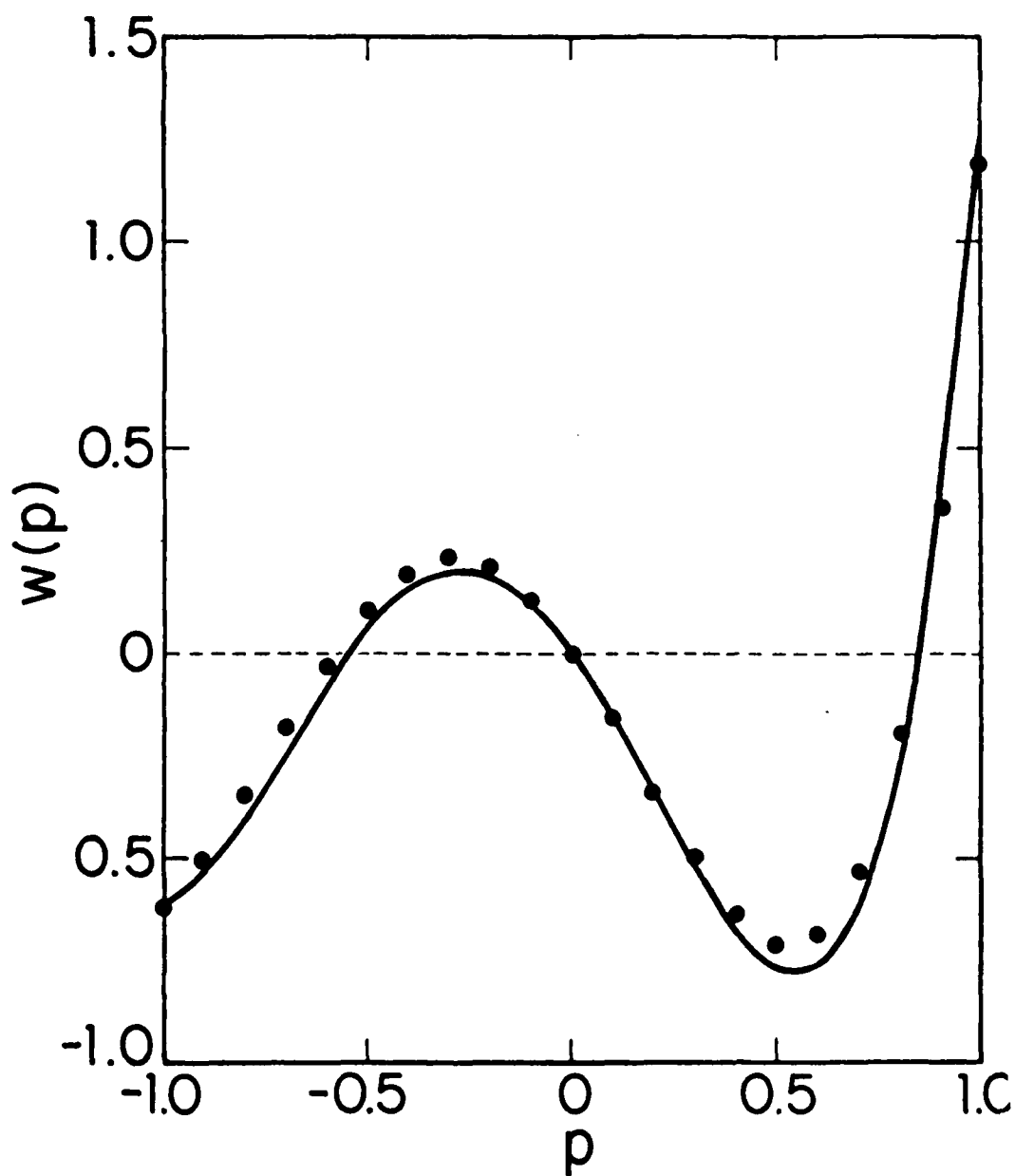
- Fig. 1 True wavefront (solid line), reconstructed wavefront realization (solid circles):  $E = .078441$ ,  $M = 31$ , 5% noise in  $\hat{\tau}$ .
- Fig. 2 True wavefront (solid line), reconstructed wavefront realization (solid circles):  $E = .078154$ ,  $M = 31$ , 5% noise in  $\hat{\tau}$ .
- Fig. 3 True wavefront (solid line), reconstructed wavefront realization (solid circles):  $E = .089435$ ,  $M = 31$ , 5% noise in  $\hat{\tau}$ .
- Fig. 4 True wavefront (solid line), reconstructed wavefront realization (solid circles):  $E = .075563$ ,  $M = 21$ , 5% noise in  $\hat{\tau}$ .
- Fig. 5 True wavefront (solid line), reconstructed wavefront realization (solid circles):  $E = .327351$ ,  $M = 21$ , 5% noise in  $\hat{\tau}$ .
- Fig. 6 True wavefront (solid line), reconstructed wavefront realization (solid circles):  $E = .097244$ ,  $M = 31$ , 10% noise in  $\hat{\tau}$ .
- Fig. 7 True wavefront (solid line), reconstructed wavefront realization (solid circles):  $E = .106523$ ,  $M = 31$ , 10% noise in  $\hat{\tau}$ .



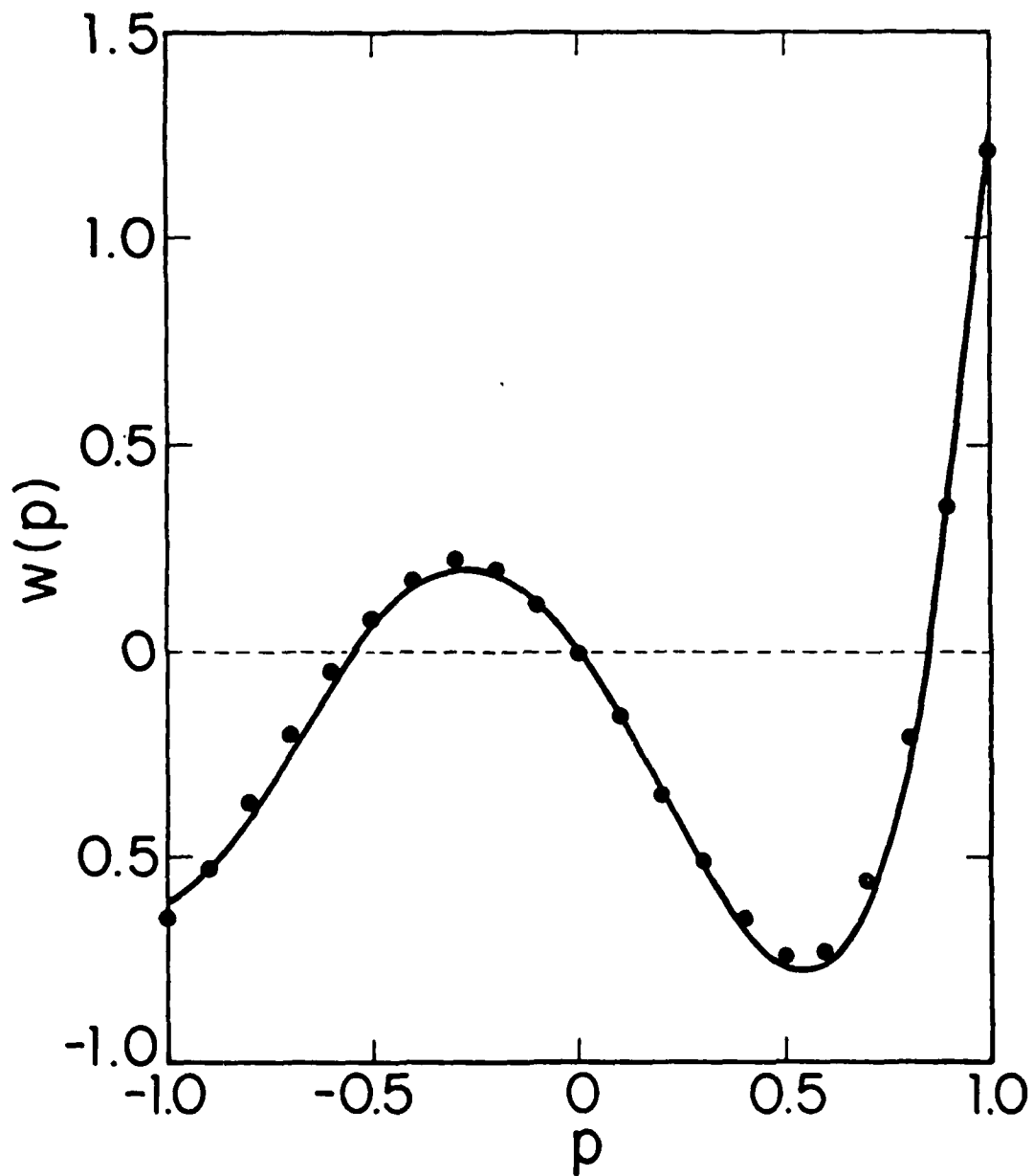
True Wavefront (Solid Line), Reconstructed Wavefront  
Realization (Solid Circles):  $E = .078441$ ,  $M = 31$ ,  
5% Noise in  $\hat{\tau}$



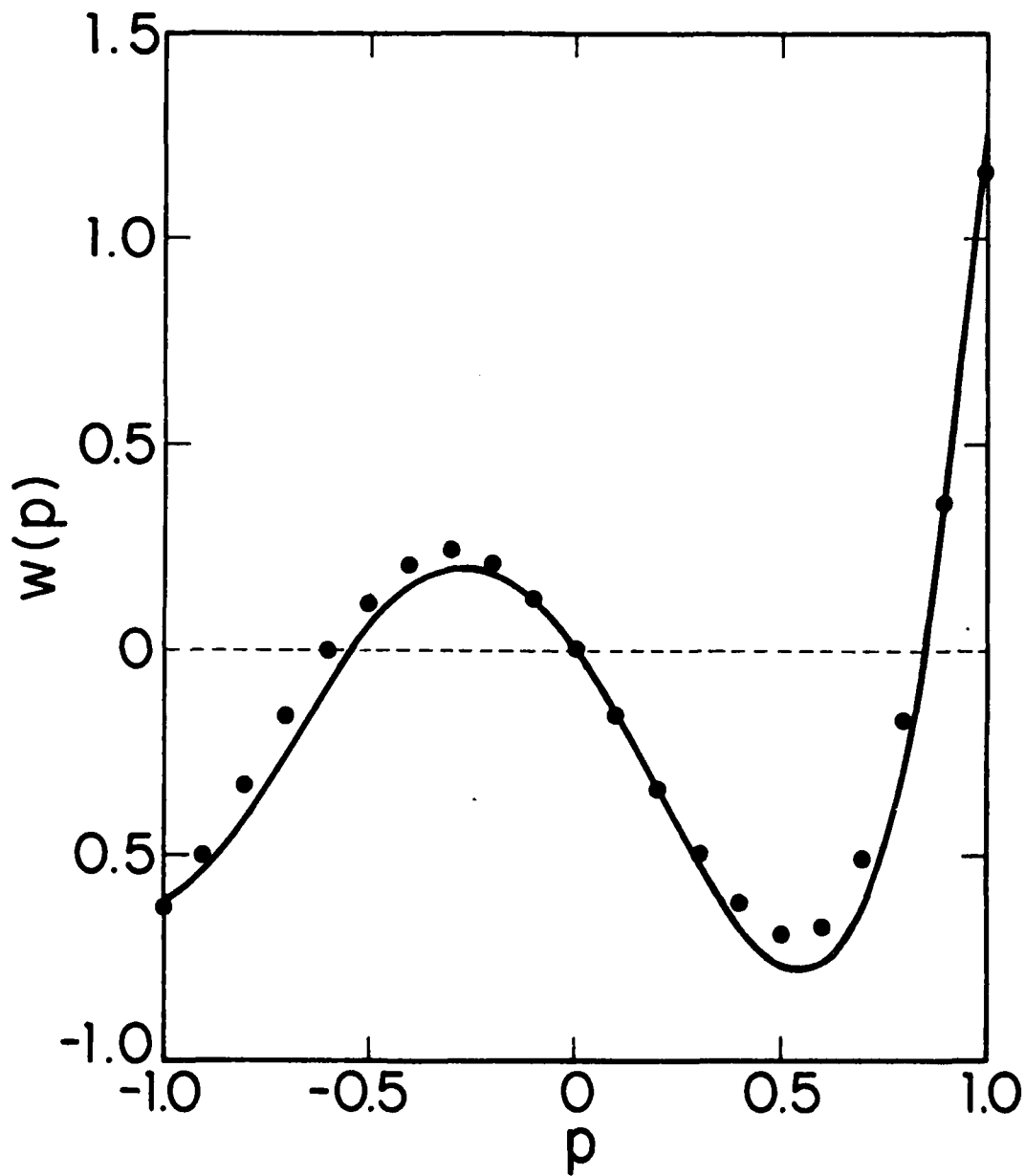
True Wavefront (Solid Line), Reconstructed Wavefront Realization (Solid Circles):  $E = .078154$ ,  $M = 31$ , 5% noise in  $\hat{\tau}$ .



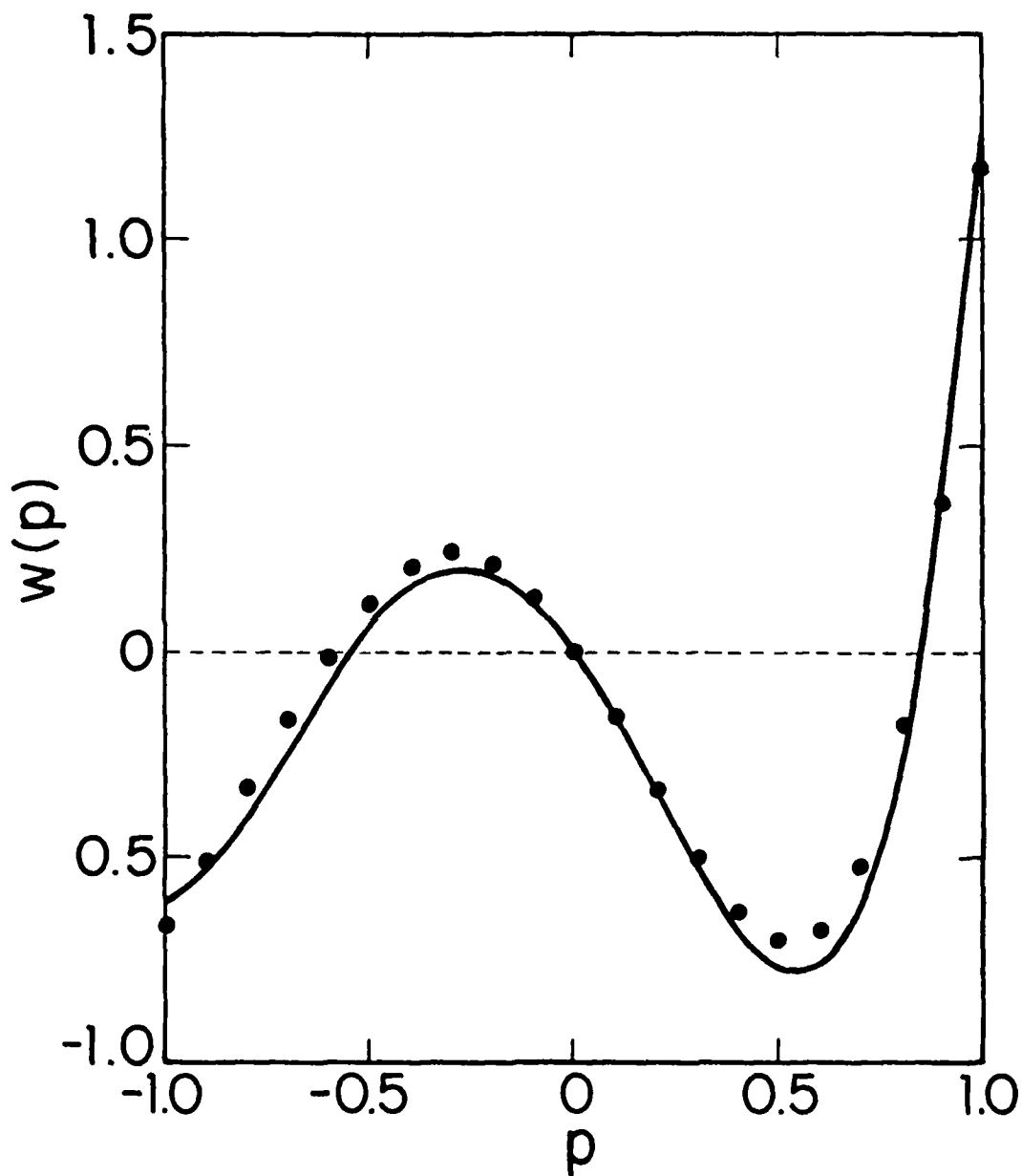
True Wavefront (Solid Line), Reconstructed Wavefront Realization (Solid Circles):  $E = .089435$ ,  $M = 31$ , 5% Noise in  $\hat{\tau}$ .



True Wavefront (Solid Line), Reconstructed Wavefront Realization (Solid Circles);  $E = .075563$ ,  $M = 21$ , 5% Noise in  $\hat{\tau}$ .

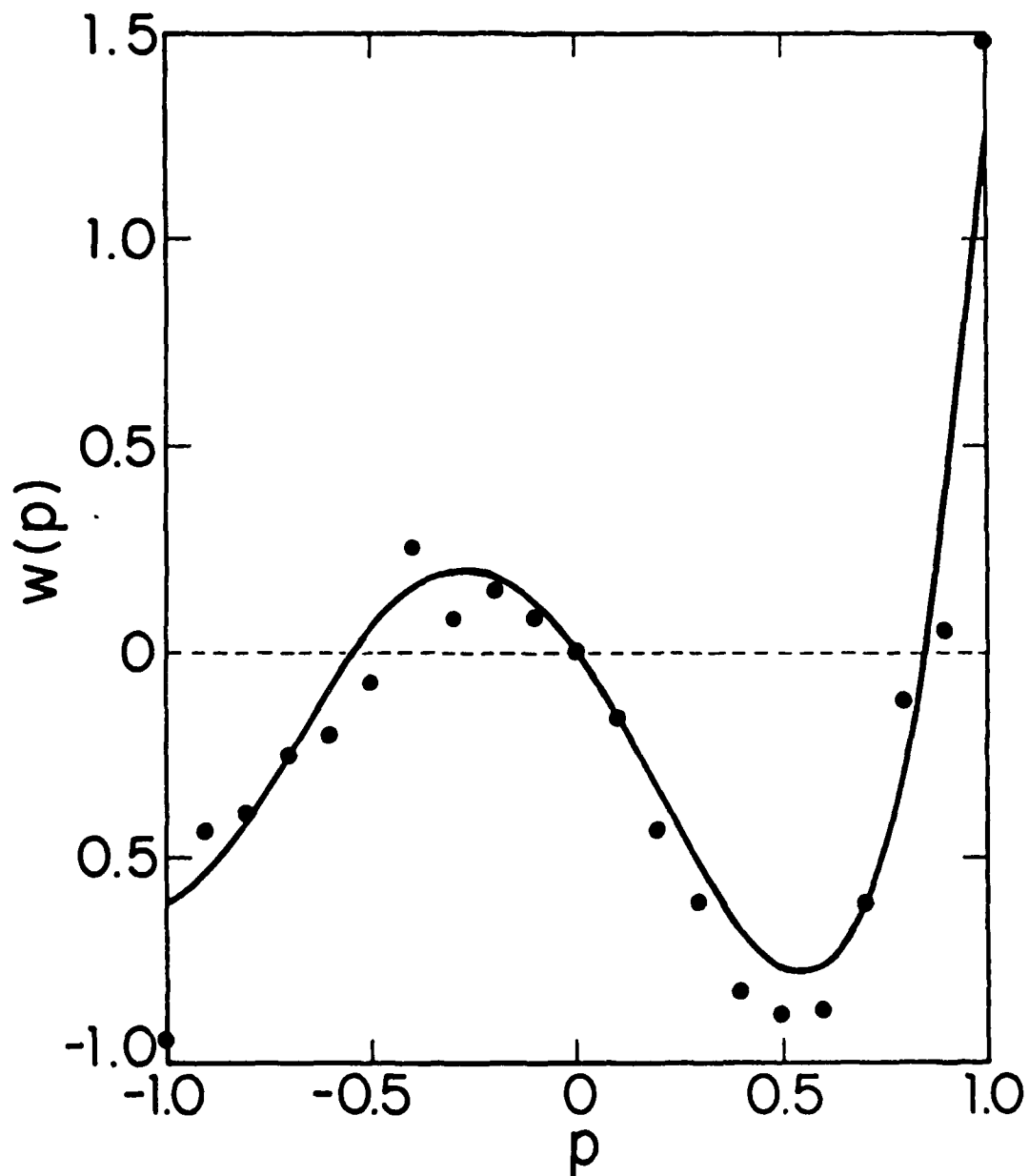


True Wavefront (Solid Line), Reconstructed Wavefront Realization (Solid Circles):  $E = .327351$ ,  $M = 21$ , 5% Noise in  $\hat{r}$ .



True Wavefront (Solid Line), Reconstructed Wavefront Realization (Solid Circles):  $E = .097244$ ,  $M = 31$ , 10% Noise in  $\hat{r}$ .





True Wavefront (Solid Line), Reconstructed Wavefront Realization (Solid Circles):  $E = .106523$ ,  $M = 31$ , 10% Noise in  $\hat{\tau}$ .

## SECTION 2

### OPTIMUM BALANCED WAVEFRONT ABERRATIONS FOR RADIALLY SYMMETRIC AMPLITUDE DISTRIBUTIONS: GENERALIZATIONS OF ZERNIKE POLYNOMIALS

#### ABSTRACT

The Zernike aberration theory for constant amplitude circular apertures is extended to annular apertures having a Gaussian-like radial taper. Explicit expressions are obtained for the optimum balanced wavefront aberrations in terms of shifted Jacobi polynomials. Properties of the polynomials (e.g., Rodrigues formula, recurrence relations, derivatives, etc.) are investigated

## 1. INTRODUCTION

The complex diffracted amplitude in the receiving plane, given that the exit pupil is circular, is

$$a(v, \phi) = \int_0^1 \int_0^{2\pi} A_0(\rho, \theta) \exp\{ikW(\rho, \theta | x_0, y_0) + i\rho v \cos(\theta - \phi)\} d\theta \rho d\rho \quad (1.1)$$

where  $W(\rho, \theta | x_0, y_0)$  is Hamilton's mixed characteristic (wavefront aberration) function with respect to the object plane coordinates  $x_0, y_0$  and  $A_0(\rho, \theta)$  is the amplitude distribution over the exit pupil. The point spread function  $t(v, \phi)$  is given by

$$t(v, \phi) \equiv \left| \frac{a(v, \phi)}{a(0, 0)} \right|^2 \quad (1.2)$$

so that  $0 \leq t(v, \phi) \leq 1$ .

For many optical systems,  $A_0(\rho, \theta)$  is constant over the aperture. Without loss of generality we set  $A_0(\rho, \theta) = 1$ ; such systems are termed Airy systems. The Zernike polynomials play a fundamental role in the diffraction theory of aberrations of Airy systems [1-17]. It is also possible to obtain the same results by direct application of Maréchal aberration balancing theory [6, 18] although not without considerable effort.

An extension of the Zernike type theory to nonconstant aperture distributions is not without interest especially those that are radially dependent, i.e.,  $A_0(\rho, \theta) = A_0(\rho)$ . We consider the

case where  $A_0(\rho)$  is given by

$$\begin{aligned} A_0(\rho) &= (1-\rho^2)^\alpha, & \epsilon \leq \rho < 1 \\ &= 0, & 0 \leq \rho < \epsilon, \quad \alpha \geq 0. \end{aligned} \quad (1.3)$$

In other words we are considering an annular aperture of inner radius  $\epsilon$  with a Gaussian-like amplitude taper. Note that for large  $\alpha$ , we have  $(1-\rho^2)^\alpha \approx \exp(-\alpha\rho^2)$  for small  $\rho$ . The situation stated in Eq. 1.3 is precisely the one encountered in active optics using a segmented annular mirror that is illuminated by a laser beam. Two special but important cases are:

$$\begin{aligned} \text{A. annular aperture } A_0(\rho) &= 1, & \epsilon < \rho < 1 \\ &= 0, & 0 \leq \rho < \epsilon \end{aligned} \quad (1.4)$$

$$\text{B. Gaussian aperture } A_0(\rho) = (1-\rho^2)^\alpha \quad (1.5)$$

Generally speaking, small to moderate amounts of wavefront aberration take energy out of the central core of the diffraction pattern and add it to the diffraction rings. Furthermore, there is very little change in the gross characteristic width of the central core of the diffraction pattern. Apodization, however,

competes with aberration effects in case B in that it takes energy out of the diffraction rings and adds them to the central core, while simultaneously broadening the characteristic width of the central core. Case A apodization behaves in a cooperative way with the aberration effects by adding even more energy to the diffraction rings while decreasing the characteristic width of the central core. The general case is intermediate. These apodization effects must manifest themselves in the determination of the optimum balanced wavefront aberrations. Obviously the nonconstant  $A_0(\rho)$  cases lead to functions that differ from the Zernike polynomials of the usual Airy system.

The purpose of the present paper is to obtain general explicit expressions for the optimum balanced wavefront aberrations. The aberration functions corresponding to Eq. 1.3 are denoted by  $C_n^m(\rho, \epsilon, \alpha)$ . The functions corresponding to case A are denoted by  $A_n^m(\rho, \epsilon)$ , those of case B by  $B_n^m(\rho, \alpha)$ . When  $\epsilon = 0$  and  $\alpha = 0$ , these functions reduce to the usual radial Zernike polynomials  $R_n^m(\rho)$ . The method employed in this paper is a generalization of the elegant (and efficient) procedure developed by Bhatia and Wolf [4] in their classic paper on Zernike polynomials. Our basic concern is with the development of explicit expressions, orthogonality conditions, recurrence relations, etc.,

and not with the diffraction imagery associated with the polynomials. However, Sec. 5 is devoted to formulae for the Strehl criterion. Evaluation of the Hankel transforms of  $C_n^m$ ,  $A_n^m$ , and  $B_n^m$  so necessary for the analytical aspects of diffraction imagery is under investigation.

## 2. DERIVATION OF $C_n^m$ POLYNOMIALS

Four conditions [5] are imposed on the radial Zernike polynomials  $R_n^m(\rho)$ :

1.  $R$ 's are orthogonal over  $(0,1)$  with weight factor unity, i.e.,

$$\int_0^1 R_n^m(\rho) R_{n'}^m(\rho) \rho d\rho = (2n+2)^{-1} \delta_{nn'}, \quad (2.1)$$

2.  $R_n^m(\rho)$  is a polynomial of degree  $n$  in  $\rho$  and its lowest term is of degree  $m$  in  $\rho$ .
3.  $R_n^m(\rho)$  is to be even or odd, the parity being the same as that of  $n$ , this means that  $n-m$  is always an even integer.
4.  $R_n^m(\rho)$  is normalized,  $R_n^m(1) \equiv 1$  for all  $n$  and  $m$ .

Condition 4 is a corollary to Condition 1.

We require that  $C_n^m(\rho)$  satisfy a modified version of these conditions. Condition 1 now becomes

$$\int_{\epsilon}^1 C_n^m(\rho)(1-\rho^2)^{\alpha} C_{n'}^m(\rho) \rho d\rho = h_n^m(\epsilon, \alpha) \delta_{nn'}, \quad (2.2)$$

where the constant  $h_n^m(\epsilon, \alpha)$  will be evaluated shortly. Conditions 2 and 3 are unchanged, while Condition 4 is modified slightly to read  $C_n^m(1) \equiv 1$  for all  $n$  and  $m$  and for  $0 \leq \epsilon < 1$ ,  $\alpha \geq 0$ .

Condition 2 is the crucial one in that the Zernike polynomials  $R_n^m$ , although orthogonal, do not form a complete set [21,22]. The procedure employed by Bhatia and Wolf [4] (and the standard procedure in such cases as the associated Legendre polynomials), is to factor out  $\rho^m$  so that the remaining polynomial is of degree  $(n-m)$ . This polynomial is orthogonal with respect to the nonnegative weight factor  $\rho^m$ . Consequently, by standard theorems in the theory of classical orthogonal polynomials [22, 23], this polynomial set is complete. Our procedure is a generalization of this.

We factor  $C_n^m(\rho)$  into two polynomials

$$C_n^m(\rho, \epsilon, \alpha) = N_n^m(\epsilon, \alpha) c_m(\rho, \epsilon) p_{n-m}(\rho, \epsilon, \alpha) \quad , \quad (2.3)$$

where the subscript denotes the degree of the polynomial.  $N_n^m$  is a normalization constant. The p-polynomials will now form a complete orthogonal set with respect to the weight factor  $[c_m(\rho, \epsilon)]^2$  over the interval  $(\epsilon, 1)$ , provided that  $c_m(\rho, \epsilon) \geq 0$  over the same interval. In fact, we will set

$$c_m(\rho, \epsilon) = \left( \frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right)^{m/2} \quad (2.4)$$

for reasons to be apparent shortly. To determine the p-polynomial, we employ the known fact [21,22] that if the weight factor  $c_m(\rho, \epsilon)$  is of the form given in Eq. 2.4, then  $p_{n-m}$  must be a



scaled version of the Jacobi polynomials  $P_k^{(\alpha, \beta)}(g)$  as defined in [20,21]. In point of fact

$$p_{n-m}(\rho, \epsilon, \alpha) \equiv P_{(n-m)/2}^{(\alpha, m)}(g) \quad , \quad (2.5)$$

where

$$g \equiv 2 \left( \frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right) - 1 \quad . \quad (2.6)$$

It remains to determine the normalization constant  $N_n^m(\epsilon, \alpha)$ . The Jacobi polynomials satisfy the condition

$$P_k^{(\alpha, \beta)}(1) = \binom{k+\alpha}{k} \quad (2.7)$$

independent of  $\beta$ . Since  $c_m(1, \epsilon, \alpha) = 1$ , it follows that

$$[N_n^m(\epsilon, \alpha)]^{-1} = \binom{\frac{n-m+\alpha}{2}}{\frac{n-m}{2}} \quad . \quad (2.8)$$

Note that if  $\alpha = 0$ , then  $N_n^m = 1$  independent of  $\epsilon$ .

Putting all these components together, we have

$$C_n^m(\rho, \epsilon, \alpha) = N_n^m(\epsilon, \alpha) \left( \frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right)^{m/2} P_{(n-m)/2}^{(\alpha, m)} \left[ 2 \left( \frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right) - 1 \right] \quad (2.9)$$

as the sought-for expression.

When  $\epsilon = \alpha = 0$ , we reduce to the usual Zernike radial polynomial

$$C_n^m(\rho, 0, 0) \equiv R_n^m(\rho) = \rho^m P_{(n-m)/2}^{(0, m)} [2\rho^2 - 1] \quad (2.10)$$

as first noted by Bhatia and Wolf [4]. Bear in mind that they used the old  $G$  notation for the shifted Jacobi polynomials.

We wish to point out that Tatian [19] had previously considered the problem of optimum balanced aberrations for the annular aperture, however, he does not derive any explicit expressions. Arimoto [20] considered the case somewhat analogous to our case B.

### 3. PROPERTIES OF $C_n^m$ POLYNOMIALS

Since  $C_n^m$  is proportional to a shifted Jacobi polynomial, we can use its properties as listed in [21,22] to derive properties of  $C_n^m$ . For typographic convenience, we omit the explicit dependence on  $C_n^m$  of  $\epsilon$  and  $\alpha$  and write  $C_n^m(\rho) = C_n^m(\rho, \epsilon, \alpha)$ .

A very useful finite series representation of  $C_n^m$  can be obtained from Eq. 4.3.2 of [21], it is

$$C_n^m(\rho) = N_n^m \left( \frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right)^{m/2} \sum_{v=0}^{(n-m)/2} \frac{\left( \frac{n-m+2\alpha}{2} \right)! \left( \frac{n+m}{2} \right)!}{v!(v+\alpha)! \left( \frac{n-m-2v}{2} \right)! \left( \frac{n+m-2v}{2} \right)!} \cdot \left( \frac{\rho^2 - 1}{1 - \epsilon^2} \right)^v \left( \frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right)^{(n-m-2v)/2} \quad (3.1)$$

Explicit expressions for the lower order  $C_n^m(\rho)$  are given in Table 1.

A Rodrigues formula for  $C_n^m$  follows from Eq. 4.3.1 of Szegő by appropriate change of variable. The final result is

$$C_n^m(\rho) = \frac{N_n^m(-1)^{(n-m)/2}}{\left[ \frac{1}{2}(n-m) \right]! \left( 1 - \epsilon^2 \right)^{(n-m)/2} (\rho^2 - \epsilon^2)^{m/2} (1 - \rho^2)^\alpha} \cdot \left( \frac{d}{d\rho^2} \right)^{(n-m)/2} \left[ (\rho^2 - \epsilon^2)^{(n+m)/2} (1 - \rho^2)^{n-m+2\alpha/2} \right] \quad (3.2)$$

A recurrence relation for fixed  $m$  and variable  $n$  can also be derived from the basic expression, Eq. 4.5.1, in Szegő. Straightforward manipulations yield

$$\begin{aligned} & \frac{1}{2}(n-m)(n+m+2\alpha)(n+\alpha-2) C_n^m(\rho) \\ &= \left[ (n+\alpha)(n+\alpha-1)(n+\alpha-2) \rho + (n+\alpha-1)(\alpha^2-m^2) \right] C_{n-2}^m(\rho) \\ & \quad - \frac{1}{2} (n+\alpha)(n-m+2\alpha-2)(n+m-2) C_{n-4}^m(\rho) \end{aligned} \quad (3.3)$$

for  $n-m \geq 2$  and  $\rho$  is given by Eq. 2.6. The initial polynomials are

$$\begin{aligned} C_n^m(\rho) &= \left( \frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right)^{m/2} \\ C_{m+2}^m(\rho) &= (1+\alpha)^{-1} (1-\epsilon^2)^{-(m+2)/2} (\rho^2 - \epsilon^2)^{m/2} [(m+2+\alpha)\rho^2 - (1+\alpha)\epsilon^2 - (m+1)] \end{aligned} \quad (3.4)$$

When  $\epsilon = \alpha = 0$ , Eq. 3.3 reduces to the recurrence relation for Zernike polynomials given in Myrick [9] and in Kintner [12].

Noll [14] has pointed out the usefulness of the derivative of the Zernike polynomials for certain applications. Both he and Kintner [10] have developed such recurrence relations. We can obtain one such relation for  $C_n^m$  in the following fashion. Differentiate both sides of Eq. 2.8 with respect to  $\rho$ , the result is

$$\begin{aligned} \frac{d}{d\rho} C_n^m(\rho) &= m\rho(\rho^2-\epsilon^2)^{-1} C_n^m(\rho) \\ &+ N_n^m \left( \frac{\rho^2-\epsilon^2}{1-\epsilon^2} \right)^{m/2} \frac{4\rho}{(1-\epsilon^2)} \frac{d}{dg} P_{(n-m)/2}^{(\alpha,m)}(g). \end{aligned} \quad (3.5)$$

The derivative (with respect to  $g$ ) on the right-hand side can be expressed in terms of the function itself by using Eq. 4.5.7 of [21]. Upon combining all terms

$$\begin{aligned} &(n+\alpha)\rho^{-1}(\rho^2-\epsilon^2)(1-\rho^2) \frac{d}{d\rho} C_n^m(\rho) \\ &= \left\{ m(n+\alpha)(1-\rho^2) - \frac{1}{2} (n-m)[(n+\alpha)g + m - \alpha] \right\} C_n^m(\rho) \\ &+ \frac{1}{2} (n+m)(n-m+2\alpha) \binom{n-2+\alpha}{n-2} \binom{n+\alpha}{n}^{-1} C_{n-2}^m(\rho). \end{aligned} \quad (3.6)$$

This formula expresses the derivative directly in terms of two polynomials of adjacent degree  $n$  and fixed order  $m$ . When  $\epsilon=\alpha=0$ , this reduces to an expression given in [15].

We now proceed to the evaluation of the coefficient  $h_n^m(\epsilon, \alpha)$  associated with the orthogonality requirement, Eq. 2.2. Upon substituting Eq. 2.9 into Eq. 2.2 and transforming to the variable  $g$  defined in Eq. 2.6, we obtain

$$h_n^m(\epsilon, \alpha) = \frac{\binom{n}{m}^2}{2^{m+\alpha+2}} \int_{-1}^{+1} (z+g)^m (1-g)^\alpha \left[ P_{(n-m)/2}^{(\alpha, m)}(g) \right]^2 dg \quad (3.7)$$

The integral has been evaluated in Szegő, Eq. 4.3.3. The final result is

$$h_n^m(\epsilon, \alpha) = \frac{\binom{n}{m}^2 (1-\epsilon^2)^{\alpha+1}}{2(n+\alpha+1)} \frac{\Gamma\left(\frac{n+m+2}{2}\right) \Gamma\left(\frac{n-m+2+2\alpha}{2}\right)}{\Gamma\left(\frac{n+m+2+\alpha}{2}\right) \Gamma\left(\frac{n+m+2+2\alpha}{2}\right)} \quad (3.8)$$

When  $\epsilon = \alpha = 0$ , this reduces to the usual result

$$h_n^m(0, 0) = (2n+1)^{-1} \quad (3.9)$$

When  $\alpha = 0$ ,  $\epsilon \neq 0$ , the resultant expression is simply

$$h_n^m(\epsilon, 0) \equiv L_n(\epsilon) = \frac{(1-\epsilon^2)}{2n+2} \quad (3.10)$$

#### 4. STREHL CRITERION

Since the aberration polynomials  $C_n^m$  are already optimum balanced, the maximum intensity (Strehl criterion) of the point spread function is at  $v = 0$  irrespective of  $\phi$ . Thus, we have

$$t(0, \phi) = \left[ \frac{\alpha+1}{2\pi(1-\epsilon^2)^{\alpha+1}} \right]^2 \left| \int_{\epsilon}^1 \int_0^{2\pi} (1-\rho^2)^{\alpha} \exp\{ikW(\rho, \theta)\} d\theta \rho d\rho \right|^2. \quad (4.1)$$

Provided that  $W(\rho, \theta)$  is small, we can expand the exponential and retain only the first three terms

$$\exp\{ikW\} \approx 1 + ikW - \frac{1}{2} k^2 W^2 + \dots \quad (4.2)$$

The term linear in  $W$  will vanish upon integration leaving

$$\begin{aligned} t(0, \phi) &\approx \left\{ 1 - \frac{k^2(\alpha+1)}{4\pi(1-\epsilon^2)^{\alpha+1}} \int_{\epsilon}^1 \int_0^{2\pi} (1-\rho^2)^{\alpha} W^2(\rho, \theta) d\theta \rho d\rho \right\}^2 \\ &\approx 1 - \frac{k^2(\alpha+1)}{2\pi(1-\epsilon^2)^{\alpha+1}} \int_{\epsilon}^1 \int_0^{2\pi} (1-\rho^2)^{\alpha} W^2(\rho, \theta) d\theta \rho d\rho \quad (4.3) \end{aligned}$$

The expansion of  $W$  at a *fixed* object point  $x_0, y_0$  is

$$\begin{aligned} W(\rho, \theta | x_0, y_0) &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} [c_{mn}(x_0, y_0) \cos m\theta \\ &\quad + s_{mn}(x_0, y_0) \sin m\theta] C_n^m(\rho) \quad (4.4) \end{aligned}$$

where  $c_{mn}(x_0, y_0)$  and  $s_{mn}(x_0, y_0)$  depend on the object point. The restrictions on  $n$  and  $m$  are that  $n \geq m$  and  $(n-m)$  is an even integer. In the special case of rotational symmetry,  $s_{mn} \equiv 0$ .

The series in Eq. 4.4 is now substituted into the integral. The orthogonality relations for the trigonometric functions and for the  $C_n^m$  polynomials allow us to obtain

$$t(0, \phi) \approx 1 - \frac{k^2(\alpha+1)}{\pi(1-\epsilon^2)^{\alpha+1}} \sum'_{nm} (c_{nm}^2 + s_{nm}^2) h_n^m(\epsilon, \alpha) \quad (4.5)$$

as the final expression for the Strehl criterion. The prime on the summation sign indicates that the terms for which  $m = 0$  are provided with a factor of one-half. When  $\epsilon = \alpha = 0$ , this reduces to

$$t(0, \phi) \approx 1 - \frac{k^2}{\pi} \sum'_{nm} (c_{nm}^2 + s_{nm}^2) (2n+2)^{-1} \quad (4.6)$$



## REFERENCES

1. F. Zernike, Beugungstheorie Des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode, *Physica*, 1, 1934, 689.
2. F. Zernike and H.C. Brinkman, Hypersphärische Funktionen und die in sphärischen Berischen Orthogonalen Polynome, *Neder. Akad. Wetensch. Indag. Math.*, 38, 1935, 161.
3. B.R.A. Nijboer, The diffraction theory of aberrations, *Physica*, 10, 1943, 679; 13, 1947, 605.
4. A.B. Bhatia and E. Wolf, On the circle polynomials of Zernike and related orthogonal sets, *Proc. Camb. Phil. Soc.*, 50, 1954, 40.
5. E.H. Linfoot, *Recent Advances in Optics* (Oxford Univ. Press, Oxford, 1955), pp. 48-66.
6. M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1964).
7. P.W. Hawkes, The diffraction theory of the aberrations of stigmatic orthomorphic optical or electron optical systems containing toric lenses or quadrupoles, *Opt. Acta*, 11, 1964, 237.

8. R. Barakat and L. Riseberg, Diffraction theory of the aberrations of a slit aperture, *J. Opt. Soc. Am.*, 55, 1965, 878.
9. D.R. Myrick, A generalization of the radial polynomials of F. Zernike, *J. SIAM Appl. Math.*, 14, 1966, 476.
10. S.N. Bezdid'ko, The use of Zernike polynomials in optics, *Sov. J. Opt. Tech.*, 41, 1974, 425.
11. E.C. Kintner, Some comments on the use of Zernike polynomials in optics, *Opt. Comm.*, 18, 1976, 235.
12. E.C. Kintner, A recurrence relation for calculating the Zernike polynomials, *Opt. Acta*, 23, 1976, 499.
13. E.C. Kintner and R.M. Sillitto, A new "analytic" method for computing the optical transfer function, *Opt. Acta*, 23, 1976, 607.
14. R.J. Noll, Zernike polynomials and atmospheric turbulence, *J. Opt. Soc. Am.*, 66, 1976, 207.
15. E.C. Kintner, On the mathematical properties of the Zernike polynomials, *Opt. Acta*, 23, 1976, 679.
16. W.J. Tango, The circle polynomials of Zernike and their application in optics, *Appl. Phys.*, 13, 1977, 327.

17. R.M. Sillitto, Diffraction of uniform and Gaussian beams: an application of Zernike polynomials, *Optik*, 48, 1977, 271.
18. A. Marechal, *Thesis* (University of Paris, 1948).
19. B. Tatian, Aberration balancing in rotationally symmetric lenses, *J. Opt. Soc. Am.*, 64, 1974, 1083.
20. A. Arimoto, Aberration expansion and evaluation of the quasi-gaussian beam by a set of orthogonal functions, *J. Opt. Soc. Am.*, 64, 1974, 850.
21. G. Szegő, *Orthogonal Polynomials* (Providence, Amer. Math. Soc., 1939).
22. E.D. Rainville, *Special Functions* (New York, MacMillan, 1960).
23. G. Freud, *Orthogonale Polynome* (Basel, Birkhauser-Verlag, 1969).

TABLE 1.  $C_n^m(\rho)$  for  $m, n \leq 6$ .

$$C_0^0 = 1$$

$$C_1^1 = (1-\epsilon^2)^{-1/2}(\rho^2-\epsilon^2)^{1/2}$$

$$C_2^0 = (1+\alpha)^{-1}(1-\epsilon^2)^{-1}[(2+\alpha)\rho^2-(1+\alpha)\epsilon^2-1]$$

$$C_2^2 = (1-\epsilon^2)^{-1}(\rho^2-\epsilon^2)$$

$$C_3^1 = (1+\alpha)^{-1}(1-\epsilon^2)^{-3/2}(\rho^2-\epsilon^2)^{1/2}[(3+\alpha)\rho^2-(1+\alpha)\epsilon^2-2]$$

$$C_3^3 = (1-\epsilon^2)^{-3/2}(\rho^2-\epsilon^2)^{3/2}$$

$$C_4^0 = (1+\alpha)^{-1}(2+\alpha)^{-1}(1-\epsilon^2)^{-2}\{(3+\alpha)(4+\alpha)\rho^4-2(3+\alpha)[2+(2+\alpha)\epsilon^2]\rho^2 \\ + [(1+\alpha)(2+\alpha)\epsilon^4+4(2+\alpha)\epsilon^2+2]\}$$

$$C_4^2 = (1+\alpha)^{-1}(1-\epsilon^2)^{-2}(\rho^2-\epsilon^2)[(4+\alpha)\rho^2-(1+\alpha)\epsilon^2-3]$$

$$C_4^4 = (1-\epsilon^2)^{-2}(\rho^2-\epsilon^2)^2$$

$$C_5^1 = (1+\alpha)^{-1}(2+\alpha)^{-1}(1-\epsilon^2)^{-5/2}(\rho^2-\epsilon^2)^{1/2}\{(4+\alpha)(5+\alpha)\rho^4 \\ - 2(4+\alpha)[3+(2+\alpha)\epsilon^2]\rho^2 \\ + [6+6(2+\alpha)\epsilon^2+(1+\alpha)(2+\alpha)\epsilon^2]\}$$

$$C_5^3 = (1+\alpha)^{-1}(1-\epsilon^2)^{-5/2}(\rho^2-\epsilon^2)^{3/2}[(5+\alpha)\rho^2-4-(1+\alpha)\epsilon^2]$$

$$C_5^5 = (1-\epsilon^2)^{-5/2}(\rho^2-\epsilon^2)^{5/2}$$

TABLE 1. (Cont.)

$$\begin{aligned}
 C_6^0 &= (1+\alpha)^{-1}(2+\alpha)^{-1}(3+\alpha)^{-1}(1-\epsilon^2)^{-3}\{(4+\alpha)(5+\alpha)(6+\alpha)\rho^6 \\
 &\quad - 3(4+\alpha)(5+\alpha)[3+(3+\alpha)\epsilon^2]\rho^4 \\
 &\quad + 3(4+\alpha)[6+6(3+\alpha)\epsilon^2+(2+\alpha)(3+\alpha)\epsilon^4]\rho^2 \\
 &\quad - [6+18(3+\alpha)\epsilon^2+9(2+\alpha)(3+\alpha)\epsilon^4+(1+\alpha)(2+\alpha)(3+\alpha)\epsilon^6]\}
 \end{aligned}$$

$$\begin{aligned}
 C_6^2 &= (1+\alpha)^{-1}(2+\alpha)^{-1}(1-\epsilon^2)^{-3}(\rho^2-\epsilon^2)\{(5+\alpha)(6+\alpha)\rho^4 \\
 &\quad - 2(5+\alpha)[4+(2+\alpha)\epsilon^2]\rho^2 \\
 &\quad + [12+8(2+\alpha)\epsilon^2+(1+\alpha)(2+\alpha)\epsilon^4]\}
 \end{aligned}$$

$$C_6^4 = (1+\alpha)^{-1}(1-\epsilon^2)^{-3}(\rho^2-\epsilon^2)^2\{(6+\alpha)\rho^6-[5+(1+\alpha)\epsilon^2]\}$$

$$C_6^6 = (1-\epsilon^2)^{-3}(\rho^2-\epsilon^2)^3$$

### SECTION 3

#### APPLICATION OF FILTERED SINGULAR VALUE DECOMPOSITION TO WAVEFRONT DECONVOLUTION

The purpose of this note is to outline a solution of the wavefront deconvolution problem using the method of filtered singular value decomposition taking direct account of the fact that noisy measurements are involved.

The basic equation is

$$\phi_i(r) = \sum_{n=1}^N B_{in}(r) z_n(r) \quad (1)$$

where

$\phi_i(r)$  = ith measured wavefront

$z_n(r)$  = phase aberration function for nth optical element

$B_{in}(r)$  = influence function connecting ith wavefront to nth optical element

We are given  $\phi_i$  and all the other data (via noisy measurements) and are required to determine  $z_n$ .

It is convenient to rewrite Eq. 1 in matrix form, so that in an obvious notation we have

$$\hat{\phi} = \hat{B} \hat{z} \quad (2)$$

where

$\hat{B}$  is size  $M \times N$  ( $M$  rows,  $N$  columns)

$\hat{z}$  is size  $N \times 1$  ( $N$  rows, 1 column)

$\hat{\phi}$  is size  $M \times 1$  ( $M$  rows, 1 column) .

For our situation  $M \geq N$ , with the most likely case being  $M = N$  (i.e., number of observations = number of unknowns).

The fundamental difficulty with ill-posed problems is the lack of sufficient information from response measurements to infer the correct solution. This is reflected (mathematically) in the fact that the system matrix  $\hat{B}$  tends to be underdetermined (rank deficient) even if it is formally overdetermined (more rows than columns). Our approach then, is to augment the data provided by the instrument with any additional knowledge of the nature of the quantity being measured in order to make the computed solution at least physically meaningful and possibly even correct. Mathematically this amounts to building up the rank of the matrix  $\hat{B}$ , or reducing the solution space so as to yield a unique solution which satisfies all constraints known to hold *a priori*. It is tempting to utilize naive least squares to "solve" this problem, i.e.,

$$\hat{B}^+ \hat{B} \hat{z} = \hat{B}^+ \hat{\phi} \quad (3)$$

where  $\hat{B}^+$  is the transpose of  $\hat{B}$ . Now  $(\hat{B}^+ \hat{B})$  is symmetric and we can formally invert to obtain  $\hat{z}$  in the form

$$\hat{z} = (\hat{B}^+ \hat{B})^{-1} \hat{B}^+ \hat{\phi} \quad . \quad (4)$$

If:

1. System matrix  $\hat{B}$  and data matrix  $\hat{\phi}$  are exact (i.e., no uncertainty in  $\hat{B}$  and  $\hat{\phi}$ ),
2.  $\hat{B}^+ \hat{B}$  is of full rank,
3. Precision of the arithmetic of the computer is such that  $\hat{B}^+ \hat{B}$  can be formed and stored exactly

then the solution  $\hat{z}$  can be obtained from  $(\hat{B}^+ \hat{B})^{-1} \hat{B}^+ \hat{\phi}$ . Unfortunately these three conditions are not to be encountered in the deconvolution problem, in that conditions 1 and 2 are not satisfied.

The difficulty is that we do not know the rank of  $\hat{B}$  and until we determine it, we cannot invert the matrix equation. The only way that we can determine the rank of  $\hat{B}$  is to use the method of singular value decomposition and thus determine the rank of  $\hat{B}$  during computation. This is not to say that there are not other methods to accomplish this but they are generally not to be trusted (e.g., ridge-regression).

We must realize that the deconvolution problem is extremely complicated because *both*  $\hat{B}$  and  $\hat{\phi}$  are measured. The vast majority of inversion problems encountered in the applied sciences have the simplifying feature that the system matrix  $\hat{B}$  is known exactly.



To solve our problem by singular value decomposition we note that the rectangular matrix  $B$  can be written as the product of three matrices

$$\hat{B} = \hat{U} \hat{\Lambda} \hat{V}^+ \quad (5)$$

where

$\hat{U} = M \times M$  orthogonal matrix ( $\hat{U} \hat{U}^+ = \hat{U}^+ \hat{U} = \hat{I}_M$ )

$\hat{V} = N \times N$  orthogonal matrix ( $\hat{V} \hat{V}^+ = \hat{V}^+ \hat{V} = \hat{I}_N$ )

$\hat{\Lambda} = M \times N$  matrix with nonnegative elements on the main diagonal and zeros elsewhere.

$\hat{\Lambda}$  has the form

$$\hat{\Lambda} \equiv \begin{vmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & \\ 0 & 0 & & \ddots \\ \dots & \dots & \dots & \dots & \sigma_N \\ & & & & \hat{0} \end{vmatrix} \equiv \begin{vmatrix} \hat{D} \\ \hat{0} \end{vmatrix} \quad (6)$$

The  $\sigma$ 's are termed the singular values of  $\hat{B}$  and are the solutions of the eigenvalue problems

$$\begin{aligned} \hat{B} \hat{B}^+ \hat{u}_j &= \sigma_j \hat{u}_j, & j &= 1, 2, \dots, M \\ \hat{B}^+ \hat{B} \hat{v}_j &= \sigma_j \hat{v}_j, & j &= 1, 2, \dots, N \end{aligned} \quad (7)$$

where  $\hat{u}_j$  and  $\hat{v}_j$  are the  $j$ th column vectors of  $U$  and  $V$ . This is the *mathematical* definition of the  $\sigma$ 's but they are never

evaluated from the definition unless  $\hat{B}$  is mathematically exact (certainly not the case with which we have to contend!).

The  $\sigma$ 's can be ordered so that

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_N \geq 0 \quad . \quad (8)$$

If the rank of  $\hat{B}$  is  $k$  where  $k < N$ , then

$$\sigma_{k+1} = \sigma_{k+2} = \dots = \sigma_N = 0 \quad . \quad (9)$$

The solution to the minimal least squares problem posed by Eq. 2 can be found in the following manner. Multiply both sides of Eq. 2 by  $\hat{B}^+$  and *formally* invert to get

$$\hat{z} = (\hat{B}^+ \hat{B})^{-1} \hat{B}^+ \hat{\phi} \quad (10)$$

Substitute Eq. 5 into the right hand side

$$\begin{aligned} (\hat{B}^+ \hat{B})^{-1} \hat{B}^+ \hat{\phi} &= (\hat{V} \hat{\Lambda}^+ \hat{U}^+ \hat{U} \hat{\Lambda} \hat{V}^+)^{-1} \hat{V} \hat{\Lambda} \hat{U}^+ \hat{\phi} \\ &= (\hat{V} \hat{\Lambda}^{-1} \hat{U}^{-1} \hat{U} \hat{\Lambda}^{-1} \hat{V}^{-1} \hat{V} \hat{\Lambda} \hat{U}^+) \hat{V}^+ \hat{\Lambda} \hat{U}^+ \hat{\phi} \\ &= \hat{V} \hat{\Lambda}^{-1} \hat{U}^+ \hat{\phi} \equiv \hat{B}^+ \hat{\phi} \quad . \end{aligned} \quad (11)$$

The matrix  $\hat{B}^+ \equiv \hat{V} \hat{\Lambda}^{-1} \hat{U}^+$  is termed the *pseudoinverse* of  $\hat{B}$ . The matrix  $\hat{\Lambda}^{-1}$  is the  $N \times M$  matrix

$$\hat{A}^{-1} = \begin{vmatrix} \sigma_1^+ & & 0 & \hat{0} \\ & \sigma_2^+ & 0 & \\ 0 & 0 & \ddots & \sigma_N^+ \\ & & \hat{0} & \hat{0} \end{vmatrix} \quad (12)$$

where

$$\begin{aligned} \sigma_j^+ &= \frac{1}{\sigma_j} \quad \text{if } \sigma_j > 0 \\ &= 0 \quad \text{if } \sigma_j = 0. \end{aligned} \quad (13)$$

The solution becomes clearer if the right hand side of Eq. 9 is written out explicitly

$$\hat{z} = \sum_{j=1}^k \left( \frac{u_j^+ \hat{\phi}}{\sigma_j} \right) \hat{v}_j, \quad k \leq N \quad (14)$$

where  $\hat{u}_j$  and  $\hat{v}_j$  denote column vectors of the matrices  $U$  and  $V$ . The smaller singular value  $\sigma_i$  entering into the denominator of the terms of the expansion tend to greatly magnify any error in the data vector  $\hat{\phi}$ , resulting in a spurious solution. To alleviate this, the expansion must be cut off (in some rational fashion) before the contamination due to the small singular values enters.

One way to achieve this is to set

$$\begin{aligned} \sigma_j^+ &= \frac{1}{\sigma_j}, \quad \sigma_j > \epsilon \\ &= 0, \quad \sigma_j < \epsilon \end{aligned} \quad (15)$$

where the criterion for picking  $\epsilon$  is

$$\frac{\epsilon}{\sigma_0} \gg \text{noise} . \quad (16)$$

This approach has been employed by the author for several inversion problems. However, it is not recommended for the deconvolution problem because it is virtually impossible to know how much noise there is in the system.

Instead we go back to Eq. 12 and introduce a filtered solution

$$\hat{z} = \sum_j [f(\sigma_1) \hat{u}_1^+ \hat{\phi}] \hat{v}_1 \quad (17)$$

where  $f(\sigma_1)$  is a filter function depending on the singular values. The filter function is required to act like  $1/\sigma_1$  for large  $\sigma_1$ , approach zero for very small  $\sigma_1$ , and finally to decrease from  $1/\sigma_1$  to zero smoothly in the intermediate range. A useful candidate is

$$f(\sigma_1) = \frac{\sigma_1^N}{\sigma_1^{N+1} + k^{N+1}} \quad (18)$$

where  $N = 0, 1, 2, \dots$  and  $k$  is a positive constant. Figures 7 and 8 show the behavior of  $f(\sigma_1)$  for  $N = 1, 3$  as a function of various values of  $k$ . Previous calculations made by the author on other (but simpler!) inversion problems have indicated that  $N = 1$  is a

desirable compromise between too much smoothing and too great a sensitivity on  $k$ . Knowing the singular values  $\sigma_1$  and requiring that  $k$  be less than the maximum singular value is one way of determining "optimum  $k$ ."

The question of *uniqueness* is a serious problem when the data is noisy. One of the useful features of the present approach is that it is possible to get a quantitative measure of the uniqueness of the solution in the presence of noisy data.

Let us put the subscript  $s$  on the solution given in Eq. 11 to denote that the solution is in terms of noisy data; hence

$$\hat{z}_s = (\hat{V}\hat{\Lambda}^{-1}\hat{U}^+)_s \hat{\phi} \equiv \hat{H}_s \hat{\phi} . \quad (19)$$

To obtain "nice"  $\hat{z}_s$ , we have had to discard singular values in  $\hat{H}_s$ . The cost we have had to pay is a degradation in "resolution" of the sought for parameters. We quantify these arguments in the following manner. Multiply Eq. 19 by  $\hat{H}_s$

$$\hat{H}_s \hat{\phi} = \hat{H}_s \hat{B} \hat{z} . \quad (20)$$

The left hand side is  $\hat{z}_s$ . We can also manipulate the right hand side

$$\begin{aligned}
\hat{z}_s &= \hat{H}_s \hat{B} \hat{z} \\
&= (\hat{V} \hat{\Lambda}^{-1} \hat{U}^+) (\hat{U} \hat{\Lambda} \hat{V}^+) \hat{z} \\
&= \hat{V} \hat{\Lambda}^{-1} \hat{\Lambda} \hat{V}^+ \hat{z} \\
&= \hat{V} \hat{V}^+ \hat{z}
\end{aligned} \tag{21}$$

Now let

$$\hat{R} \equiv \hat{V} \hat{V}^+ \tag{22}$$

so that

$$\hat{z}_s = \hat{R} \hat{z} . \tag{23}$$

Thus, the degree of which  $\hat{R}$  approximates the unit matrix  $\hat{I}$  is a measure of the uniqueness of the solution.

There is also an interpretation of the  $\hat{U}$  matrix, although it is not as important as the  $\hat{V}$  matrix interpretation. Consider the basic equation again

$$\hat{\phi} = \hat{B} \hat{z} = \hat{U} \hat{\Lambda} \hat{V}^+ \hat{z} \tag{24}$$

and set

$$\hat{Z} \equiv \hat{V}^+ \hat{z} \tag{25}$$

then

$$\hat{\phi} = \hat{U} \hat{\Lambda} \hat{Z} \tag{26}$$

If we let

$$\hat{\phi} \equiv \hat{U}^+ \hat{\phi} \quad (27)$$

then

$$\hat{\phi} = \hat{\Lambda} \hat{Z} \quad (28)$$

Now when noise is present, let us multiply both sides by  $\hat{U}$

$$\hat{U} \hat{\phi} = \hat{U} \hat{U}^+ \hat{\phi} \quad (29)$$

Set

$$\hat{S} \equiv \hat{U} \hat{U}^+ \quad (30)$$

Since our data is noisy  $\hat{S}$  is not the unit matrix so that

$$\hat{U} \hat{\phi} = \hat{S} \hat{\phi} \quad .$$

The left hand side is really  $\hat{\phi}_s$ , the noisy model data, hence

$$\hat{\phi}_s = \hat{S} \hat{\phi} \quad (31)$$

This means that our model data  $\hat{\phi}_s$  will deviate more and more from  $\hat{\phi}$  the more the  $\hat{S}$  matrix deviates from the unit matrix.

Thus, both  $\hat{U}$  and  $\hat{V}$  matrices from the singular value decomposition of  $\hat{B}$  are useful in determining the robustness of the inversion calculations.

The elements of the system matrix  $\hat{B}$  as we have already point out are experimentally determined quantities. For the ill conditioned  $\hat{B}$ , we must accept the fact that the rank of the matrix

is poorly determined numerically and may well change as the matrix elements vary by very small amounts. An added advantage of singular value decomposition is that the singular values are stable to perturbations in the matrix elements in that perturbations of the matrix elements produce perturbations in the singular values of the same order of magnitude. This is certainly not the case with the corresponding eigenvalues, should they exist!

Thus far we have discussed the *strategy* of the method of singular value decomposition. The *tactics* (i.e., the actual programs, etc.) are fortunately available. Golub and Reinech have developed an ingenious method of computing the singular values of an  $M \times N$  matrix which is numerically very stable. The algorithm itself is too complicated to describe as it employs methods generally known only to specialists in numerical linear algebra.



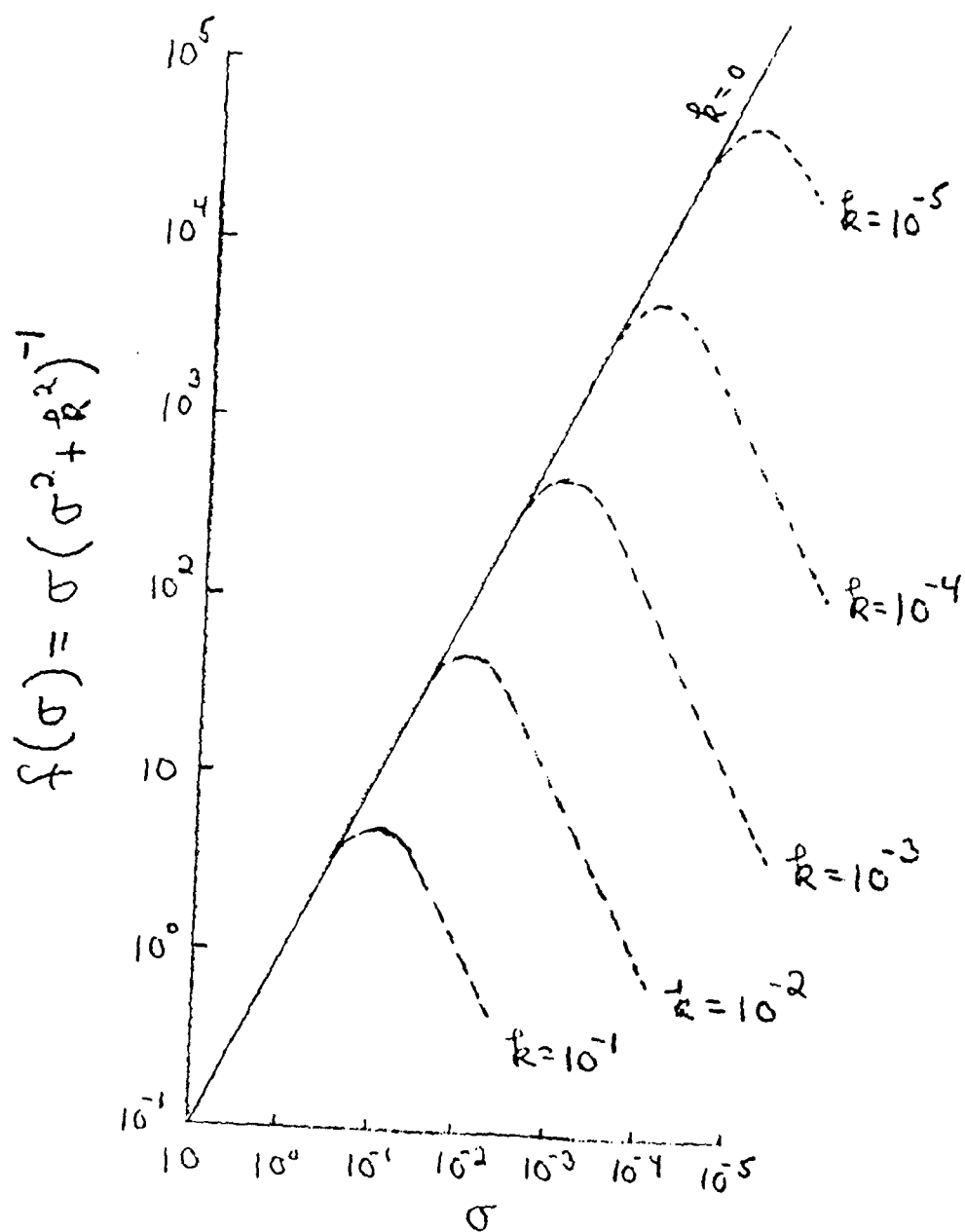


FIG 8  
Behavior of  $f(\sigma_1)$  for  $N = 1, 3$  as a Function of Various Values of  $k$

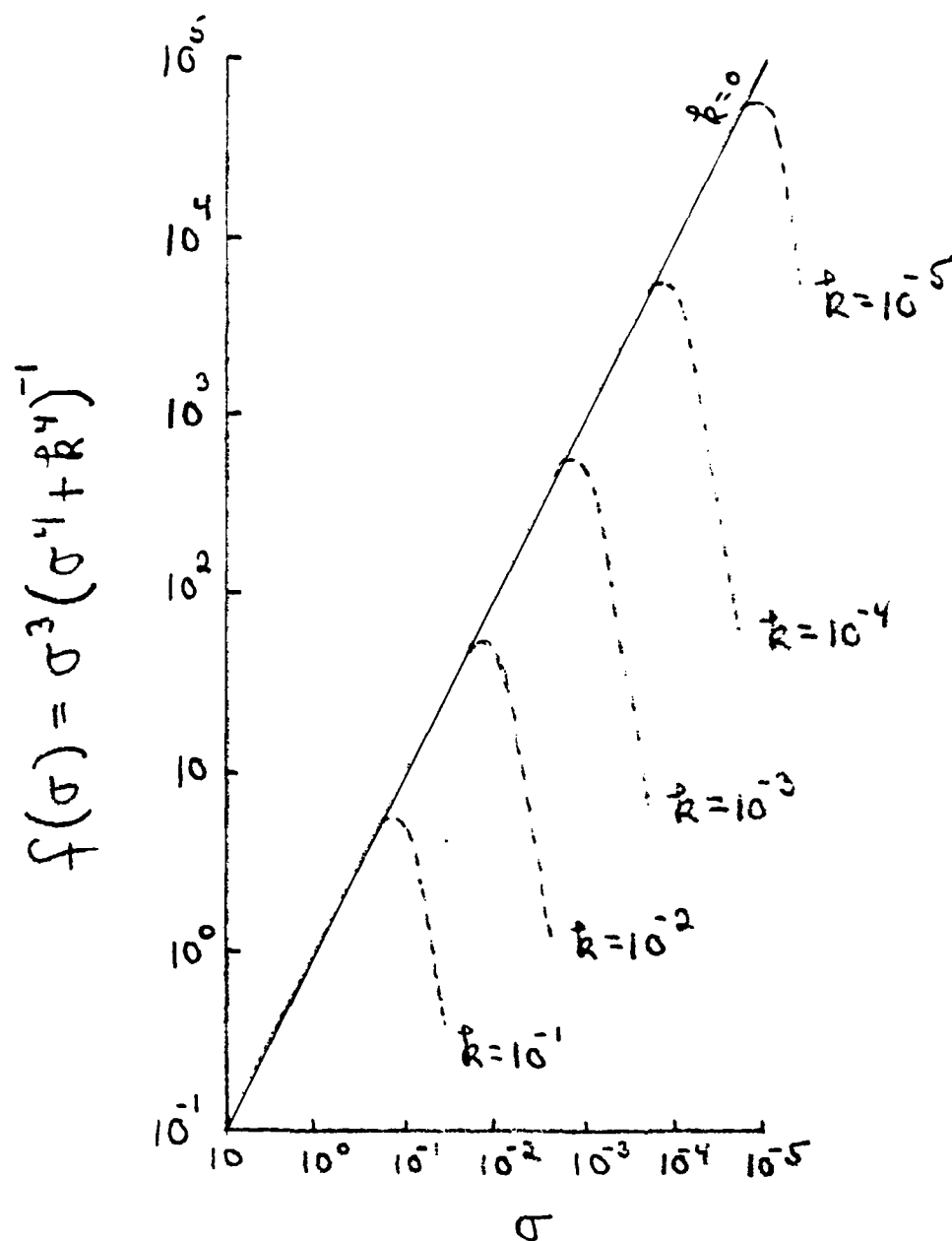


FIG 9

Behavior of  $f(\sigma_1)$  for  $N = 1, 3$  as a Function of Various Values of  $R$ .



## MISSION of *Rome Air Development Center*

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C<sup>3</sup>I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.